

ROLAP e MOLAP:
Comparação e Avaliação Prática

por

Soraia Vanessa Moura Velho

Dissertação apresentada como requisito
parcial para obtenção do grau de

Mestre em Estatística e Gestão de Informação

pelo

Instituto Superior de Estatística e Gestão da Informação

da

Universidade Nova de Lisboa

Instituto Superior de Estatística e Gestão da Informação

Universidade Nova de Lisboa

ROLAP e MOLAP:

Comparação e Avaliação Prática

Soraia Vanessa Moura Velho

Dissertação apresentada como requisito parcial para a obtenção do grau de

Mestre em Estatística e Gestão de Informação

Professor orientador:

Professor Doutor Miguel de Castro Neto

Novembro, 2011

AGRADECIMENTOS

Queria agradecer a todos os que tornaram possível a realização deste trabalho.

Aos meus colegas e amigos, em especial à Rita, Paula, Tiago, Frederico e André pela disponibilidade, ajuda e incentivo.

À minha família, em especial aos meus pais, irmão e avós por todo o apoio e compreensão.

Ao Professor Doutor Miguel Neto que se disponibilizou a orientar esta dissertação.

RESUMO

Cada vez existem mais dados e informação. No entanto, a informação só é útil para as empresas se for utilizada como sistema de apoio à decisão. Para que tenha esta utilidade podem-se criar relatórios que permitam uma visualização apenas da informação mais relevante. Estes relatórios são criados com ferramentas de reporting, mas as bases de dados online transaction processing (OLTP) são limitantes quando se utilizam estas ferramentas, por isso apareceu o online analytical processing (OLAP).

O OLAP permite uma rápida análise dos dados multidimensionais. As bases de dados OLAP podem ter várias arquitecturas, sendo as principais: relational online analytical processing (ROLAP) e multidimensional online analytical processing (MOLAP). O ROLAP armazena os dados em bases de dados relacionais, enquanto o MOLAP armazena em bases de dados multidimensionais (cubos).

O objectivo deste trabalho é comparar estas duas arquitecturas de fontes de dados a fim de descobrir se, para o propósito de reporting, existem diferenças entre elas. Para tal utilizou-se a ferramenta MicroStrategy, onde se criaram relatórios iguais utilizando como fontes as duas arquitecturas referidas. Estes relatórios foram comparados em dois aspectos: tempo que demoraram a executar e o tamanho da cache.

Estas comparações permitiram concluir que nas condições do presente estudo os relatórios ROLAP e MOLAP apresentam diferenças. Os relatórios ROLAP demoram menos tempo a serem executados, mas têm caches maiores que os MOLAP.

Palavras-chave:

Online Analytical Processing (OLAP), Relational Online Analytical Processing (ROLAP), Multidimensional Online Analytical Processing (MOLAP)

ABSTRACT

Data and information are always increasing. However this information is only useful for enterprises if used as a decision support system. For the information to be used as such, reports that display only the most relevant data can be created. These reports are created with reporting tools, but due to online transaction processing (OLTP) databases being limited when using these tools, online analytical processing (OLAP) appeared.

OLAP allows a quick analysis of multidimensional data. OLAP databases may have several architectures, the two main being relational online analytical processing (ROLAP) and multidimensional online analytical processing (MOLAP). ROLAP stores the data in relational databases, while MOLAP stores the data in multidimensional databases (cubes).

The goal of this study is to compare these two architectures of data sources to find out if there is any difference between them for reporting purposes. To achieve this, it was used MicroStrategy software where equal reports using these two architectures of data sources were created. These reports were compared through execution time and cache size.

These comparisons allowed us to conclude that in conditions of this study ROLAP and MOLAP reports are different. ROLAP reports have a shorter execution time, but they have larger cache sizes.

Keywords:

Online Analytical Processing (OLAP), Relational Online Analytical Processing (ROLAP), Multidimensional Online Analytical Processing (MOLAP)

ÍNDICE

AGRADECIMENTOS	iii
RESUMO.....	iv
ABSTRACT.....	v
ÍNDICE.....	vi
ÍNDICE DE QUADROS	viii
ÍNDICE DE FIGURAS	ix
LISTA DE SIGLAS E ABREVIATURAS	x
 1. INTRODUÇÃO	 1
1.1. Enquadramento Geral do Estudo	1
1.2. Objectivo do Estudo.....	3
1.3. <i>Rationale</i> / Lógica.....	4
1.4. Formulação das Questões.....	5
1.5. Delimitações do Estudo.....	5
 2. REVISÃO DA LITERATURA.....	 7
2.1. OLTP e OLAP	10
2.2. Importância OLAP.....	14
2.3. Tipos de OLAP	14
 3. REPORTING	 20
3.1. OLAP Reporting	20
3.2. Possíveis Abordagens para Relatórios	21
 4. METODOLOGIA	 23
4.1. Investigação-Acção.....	23
4.2. Apresentação da Empresa	24
4.3. Apresentação/Caracterização dos Dados	24
4.4. Descrição das Etapas na Criação dos Relatórios	27
4.4.1. Base de dados.....	28
4.4.2. Database instance.....	28
4.4.3. Projecto.....	29
4.4.4. Relatórios.....	29
4.5. Apresentação dos Relatórios	30

4.5.1. Relatórios 1 a 5.	31
4.5.1.1. Relatório 1.....	31
4.5.1.2. Relatório 2.....	32
4.5.1.3. Relatório 3.....	32
4.5.1.4. Relatório 4.....	34
4.5.1.5. Relatório 5.....	35
4.5.2. Relatórios 6 e 7.	36
4.5.2.1. Relatório 6.....	36
4.5.2.2. Relatório 7.....	36
4.5.3. Relatórios 8 a 10.	37
4.5.3.1. Relatório 8.....	38
4.5.3.2. Relatório 9.....	38
4.5.3.3. Relatório 10.....	39
4.5.4. Relatórios 11 e 12.	40
4.5.4.1. Relatório 11.....	40
4.5.4.2. Relatório 12.....	41
4.5.5. Relatório 13.	41
4.5.5.1. Relatório 13.....	41
4.5.6. Relatório 14.	42
4.5.6.1. Relatório 14.....	42
4.5.7. Relatórios 15 a 18.	43
4.5.7.1. Relatório 15.....	45
4.5.7.2. Relatório 16.....	46
4.5.7.3. Relatório 17.....	46
4.5.7.4. Relatório 18.....	46
4.6. Limitações ROLAP e MOLAP.....	48
5. APRESENTAÇÃO DOS RESULTADOS	50
6. DISCUSSÃO DOS RESULTADOS	57
7. CONCLUSÃO E RECOMENDAÇÕES	59
7.1. Conclusão	59
7.1. Limitações	60
APÊNDICE A - Comparação dos Resultados dos Relatórios	61
REFERÊNCIAS	66

ÍNDICE DE QUADROS

Quadro 2.1. <i>Diferenças entre OLTP e OLAP</i>	13
Quadro 3.1. <i>Diferenças entre Actividades Operacionais e Actividades Baseadas em Análise, Orientadas à Decisão</i>	20
Quadro 4.1. <i>Apresentação dos Relatórios</i>	47
Quadro 5.1. <i>Comparação dos Resultados dos Relatórios</i>	50
Quadro 7.1. <i>Resultados dos Relatórios ROLAP</i>	61
Quadro 7.2. <i>Resultados dos Relatórios MOLAP</i>	63

ÍNDICE DE FIGURAS

<i>Figura 1.1.</i> Quadrante mágico para plataformas de BI de 2011.....	2
<i>Figura 2.1.</i> Exemplo de uma tabela de factos e quatro tabelas de dimensões, tendo como modelo lógico o <i>star schema</i>	16
<i>Figura 2.2.</i> Exemplo de um cubo.....	17
<i>Figura 4.1.</i> Arquitectura ROLAP.	25
<i>Figura 4.2.</i> Arquitectura MOLAP.....	26
<i>Figura 4.3.</i> Esquema das etapas na criação dos relatórios.	27
<i>Figura 4.4.</i> Relatório 3 ROLAP no MicroStrategy.....	34
<i>Figura 4.5.</i> Relatório 7 MOLAP no MicroStrategy.....	37
<i>Figura 4.6.</i> Relatório 9 ROLAP no MicroStrategy.....	39
<i>Figura 4.7.</i> Relatório 13 MOLAP no MicroStrategy.	42
<i>Figura 4.8.</i> Exemplo de um relatório simples no MicroStrategy.	44
<i>Figura 4.9.</i> Exemplo de um relatório com um <i>derived element</i> no MicroStrategy.	45
<i>Figura 5.1.</i> Comparação do tempo de execução entre relatórios ROLAP e MOLAP.	51
<i>Figura 5.2.</i> Comparação da cache entre relatórios ROLAP e MOLAP.	52

LISTA DE SIGLAS E ABREVIATURAS

BI – *Business Intelligence*

b-on - Biblioteca do Conhecimento Online

DROLAP - *Dynamic Online Analytical Processing*

DW - *Data Warehouse*

FASMI - *Fast Analysis of Shared Multi-dimensional Information*

HOLAP - *Hybrid Online Analytical Processing*

IC - *Intelligent Cube*

KPI - *Key Performance Indicator*

MDDB - *Multidimensional Database*

MDX - *Multidimensional Expressions*

MOLAP - *Multidimensional Online Analytical Processing*

ODBC - *Open Database Connectivity*

O3LAP - *Object-Oriented Online Analytic Processing*

OLAP - *Online Analytical Processing*

OLTP - *Online Transaction Processing*

OO - *Object-Oriented*

RDBMS – *Relational Database Management System*

ROLAP - *Relational Online Analytical Processing*

SAD - *Sistemas de Apoio à Decisão*

SSAS - *SQL Server Analysis Services*

SQL - *Structured Query Language*

SSAS - *SQL Server Analysis Services*

WoS - *Web of Science*

1. INTRODUÇÃO

1.1. Enquadramento Geral do Estudo

Todas as empresas estabelecem objectivos que se propõem atingir. No entanto, para que os consigam alcançar têm de ir tomando várias decisões que irão influenciar o seu futuro e que podem até comprometer esses seus objectivos.

Para ajudar os gestores das empresas a tomarem as decisões que melhor conduzem a esses objectivos existem os SAD (sistemas de apoio à decisão). Os SAD são “sistemas de informação interactivos baseados em computadores que suportam a tomada de decisão e a gestão do conhecimento” (Asghar, 2009). Alguns destes SAD utilizam ferramentas de *reporting*, i.e., software cujo objectivo é gerar relatórios que permitam visualizar dados de diferentes fontes, para ajudar os gestores a tomarem as suas decisões, através da visualização apenas da informação relevante.

Os relatórios produzidos para ajudar os SAD podem ser realizados com bases de dados OLTP (*online transaction processing*) ou OLAP (*online analytical processing*). OLTP envolve o processamento de dados de sistemas transaccionais, lidando com um grande volume de dados, enquanto o OLAP envolve um menor volume de dados transaccionais, mas consultas à base de dados mais complexas. OLAP é também ele uma ferramenta de apoio à decisão (Datawarehouse4u.Info, 2008-2009).

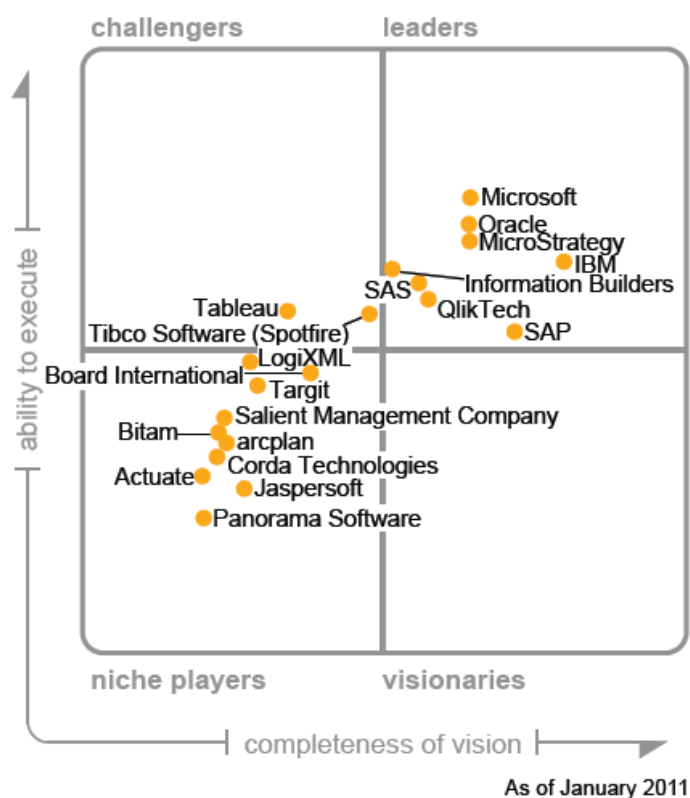
Devido à importância que o OLAP tem para os SAD e consequentemente para o desempenho das empresas, irei abordar o conceito de OLAP e os seus vários tipos.

O OLAP surgiu como uma necessidade de responder a consultas analíticas multi-dimensionais (Zhou & Chen, 2009). Existem vários tipos de OLAP, sendo o ROLAP (*relational online analytical processing*) e o MOLAP (*multidimensional online analytical processing*) os principais (Zhou & Chen, 2009), mas existe ainda o HOLAP (*hybrid OLAP*) e o *desktop OLAP* (Kazi et al., 2010).

A escolha entre o ROLAP e o MOLAP depende de muitos factores, principalmente do nível de performance de cada uma das opções, mas também é influenciada pelo software que se pretende utilizar. Cada software tem as suas

características, sendo uns mais especializados para trabalhar com ROLAP e outros com MOLAP, mas como actualmente existem vários softwares e, consequentemente, vários vendedores de software, a escolha por um deles revela-se complicada.

No entanto, para facilitar essa escolha pode-se recorrer ao Quadrante Mágico, publicado pela empresa Gartner, no qual se distribui pelos quatro quadrantes (líderes, desafiantes, visionários e nichos) os principais vendedores de software. O Quadrante Mágico para plataformas de BI (*business intelligence*) publicado no início de 2011 é apresentado na Figura 1.1.



Source: Gartner (January 2011)

Figura 1.1. Quadrante mágico para plataformas de BI de 2011.

Fonte: Feiman & MacDonald, 2011

Na Figura 1.1., a MicroStrategy encontra-se posicionada como líder, sendo este quadrante descrito como fazendo parte dele os “fornecedores de software que são significativamente fortes na amplitude e na profundidade das funcionalidades das suas plataformas de BI e que conseguem implementar soluções por todo o ambiente empresarial suportando uma ampla estratégia de BI” (Feiman & MacDonald, 2011). No

mesmo relatório onde se encontra o Quadrante Mágico, a MicroStrategy é apresentada como sendo “especializada em implementações de BI sobre grandes *data warehouses*” (Feiman & MacDonald, 2011). Refere-se também que esta plataforma é “eficiente, permitindo que um pequeno número de administradores suporte grandes projectos de BI com muitos utilizadores, com complexos requisitos de análise e relatórios e com uma grande quantidade de dados” (Feiman & MacDonald, 2011). Destaca-se ainda a sua capacidade para “criar relatórios e outros conteúdos analíticos com um elevado grau de formatação e sofisticação analítica com menos esforço e custo do que outras plataformas” (Feiman & MacDonald, 2011).

No caso da MicroStrategy, esta tem sido reconhecida como estando mais vocacionada para ROLAP (Orbit, 2009; DataWarehousingGuide.com, 2011). Mas para ser possível afirmar-se tal, tem de se comparar a ferramenta de *reporting* com os dois tipos de dados e comprovar se, de facto, o que intuitivamente se pensa está realmente correcto. Por isso, pretende-se utilizar a plataforma MicroStrategy para analisar e comparar as diferenças entre relatórios com dados provenientes de fontes ROLAP e MOLAP.

Ao longo deste estudo irei abordar esta ferramenta de *reporting* e os tipos de base de dados (ROLAP e MOLAP).

1.2. Objectivo do Estudo

Com este estudo pretendo analisar as principais diferenças, vantagens e desvantagens entre dois tipos de sistemas OLAP (ROLAP e MOLAP) e retirar algumas conclusões sobre qual o melhor ou mais adequado para o propósito de *reporting*.

Para tal irei utilizar o MicroStrategy por ser uma das plataformas líderes nas plataformas de BI.

Através da utilização da ferramenta MicroStrategy, também pretendo criar relatórios com dados provenientes de cubos e de tabelas. Ao comparar estes relatórios

espero confirmar se realmente a MicroStrategy é mais adequada para dados ROLAP em vez de dados MOLAP como sugerido por Golden Orbit (Orbit, 2009; DataWarehousingGuide.com, 2011).

A escolha deste tema surgiu na sequência de actualmente estar a trabalhar com o software MicroStrategy, o qual utilizo tanto com dados provenientes de cubos, como de tabelas.

1.3. *Rationale / Lógica*

Actualmente a quantidade de dados existente é cada vez maior. As empresas recolhem cada vez mais dados sobre o seu negócio e os seus clientes (Asghar & Osama, 2009; Hasan et al., 2000), transformando-os em informação. No entanto, só uma parte desses dados é verdadeiramente relevante e pode ajudar a decidir o futuro da empresa. Uma forma de obter apenas os dados relevantes é apresentar os principais KPIs (*key performance indicators*) em relatórios e de forma a que sejam facilmente analisados (Hasan et al., 2000). Com isto pretende-se ajudar os gestores das empresas para que consigam analisar a informação rapidamente e percebam quais as decisões a tomar e que melhor conduzam aos objectivos estabelecidos pela empresa (John, 2008). Por isso é importante para as empresas que “exista uma visão analítica dos dados, uma vez que esta pode conduzir a uma tomada de decisão eficaz” (Asghar & Osama, 2009). É por esta elevada importância da análise dos dados e por o negócio estar em constantes alterações que é necessário se estar a alterar e/ou criar novos relatórios numa empresa. Deste modo, é importante analisar e comparar os vários tipos de armazenamento de dados para que se chegue à conclusão de qual o mais adequado para este propósito (quais os motivos que podem levar a optar-se por um deles, as vantagens, as desvantagens, rapidez, performance e espaço ocupado por cada um deles) e as suas limitações.

1.4. Formulação das Questões

Com o aparecimento do termo OLAP, em 1993, apareceram também vários tipos de OLAP. Os principais são o ROLAP e MOLAP, no entanto existem ainda o HOLAP e o *desktop OLAP* (Kazi et al., 2010).

Como estes tipos de OLAP são tão diferentes, apareceram também estudos com o intuito de compará-los, principalmente o ROLAP com o MOLAP. Destas comparações surgiram várias opiniões. Uma delas defende que o ROLAP não é apropriado para o OLAP (Colliat, 1996), enquanto outras defendem que depende da quantidade de dados utilizada (Pirnau & Botezatu, 2010; Kazi et al., 2010). Houve quem também os comparasse, mas de forma a analisar a influência da aceitação de uma nova tecnologia, concluindo que “as ferramentas MOLAP fazem com que os sistemas de *data warehouse* sejam mais fáceis mas não tão úteis, enquanto as ferramentas ROLAP fazem com que o *data warehouse* seja útil mas difícil de utilizar” (Gorla, 2003).

Neste trabalho também irei comparar os dois principais tipos de OLAP: o ROLAP e o MOLAP. Para tal, espero responder às seguintes questões:

1. Quais são as vantagens, desvantagens e limitações do ROLAP e do MOLAP?
2. Existem diferenças entre relatórios desenvolvidos no MicroStrategy utilizando fontes de dados ROLAP e MOLAP?

1.5. Delimitações do Estudo

Apesar da ferramenta MicroStrategy ser considerada uma plataforma de BI líder, acaba por ter algumas especificidades. Aliás, cada ferramenta tem as suas características, podendo ser mais vocacionada para um dos tipos de OLAP. Por isso, se se for considerada a generalização das conclusões a que se chegar neste estudo relativas

à comparação de relatórios MOLAP versus ROLAP, deve-se ter em consideração que as ferramentas de *reporting* não são todas iguais e apresentam características diferentes.

2. REVISÃO DA LITERATURA

Neste capítulo irá apresentar-se o que já se investigou e escreveu sobre OLTP, OLAP e dois dos tipos de OLAP: ROLAP e MOLAP. Como é expectável, as questões de investigação abordam estes temas. A primeira questão consiste nas vantagens, desvantagens e limitações dos dois principais tipos de sistemas OLAP (que, como já se referiu anteriormente, é o ROLAP e MOLAP) e, a segunda, em descobrir se existe diferenças entre relatórios desenvolvidos no MicroStrategy, utilizando fontes de dados ROLAP e MOLAP.

A revisão da literatura foi efectuada entre Outubro de 2010 e Agosto de 2011, tendo as pesquisas sido realizadas em bases de dados disponíveis em diversos portais, como a b-on e a WoS (*Web of Science*), principalmente em inglês, por todos os tipos de documentos e restringindo-se aos anos de 1970 a 2011. Inicialmente começou por se pesquisar apenas documentos mais recentes (2000 a 2011), mas depois considerou-se alargar este intervalo. Esta decisão foi tomada porque as ideias/conceitos OLTP, OLAP e, conseqüentemente, ROLAP e MOLAP, são relativamente recentes, por isso considerou-se que um alargamento do período temporal poderia retornar algum documento importante que reportasse os seus aparecimentos.

As palavras-chave foram seleccionadas de acordo com os principais temas que irei abordar. Uma das palavras-chave é OLTP, que ajudará a perceber o aparecimento do OLAP, que é outra das palavras-chave que irá introduzir e ajudar a perceber melhor os temas que se seguem. As outras palavras-chave são ROLAP e MOLAP, pois é na comparação entre estes dois termos que o trabalho se focará. Contudo, estas duas últimas palavras-chave têm sinónimos. *Relational OLAP* e *relational databases* foram as expressões encontradas que significam o mesmo que o termo ROLAP. *Multidimensional OLAP* e *multidimensional databases* foram as expressões encontradas como sinónimos do termo MOLAP.

Dos artigos lidos durante a pesquisa, seleccionou-se os que se consideraram mais importantes e que acrescentariam informação essencial para um maior conhecimento do tema.

Com esta revisão da literatura espera-se conhecer melhor cada um dos temas e, principalmente, conhecer o que já se investigou sobre eles e quais as opiniões de quem o fez.

Este capítulo começará por abordar os temas OLTP e OLAP, o que irá permitir um conhecimento mais aprofundado sobre os seus conceitos e características. Também se irá comparar ambos os termos, à semelhança do que foi feito por Chaudhuri e Dayal, em 1997, que começaram por comparar OLTP com OLAP antes de abordarem os temas de *data warehousing* e OLAP. No ano seguinte, Bispo (1998) também os compara na sua dissertação de Mestrado, onde apresenta as três ferramentas da nova geração de Sistemas de Apoio à Decisão, sendo uma delas o OLAP. Alguns anos mais tarde, Conn (2005) compara-os com o intuito de justificar a integração dos dois sistemas e Lenz e Thalheim, em 2006, estabelecem também uma comparação entre OLTP e OLAP num artigo sobre aplicações OLAP, em que chegam a uma arquitectura para OLTP-OLAP aplicações. Em 2008, John, no seu relatório de seminário sobre OLAP, estabelece também a comparação entre ambos. Também em 2008, Bog, Krueger e Schaffner estabelecem uma comparação entre ambos quando defenderam a criação de um novo *benchmark*/referência para o OLTP e os relatórios operacionais. Passado um ano, Bog, Domschke, Mueller e Zeier (2009) propõem um sistema que combine o OLTP e OLAP e para tal começam por compará-los. Também em 2009, Lenz e Thalheim os comparam novamente quando apresentam propriedades que garantam ou contradigam a correcção dos cálculos OLAP.

A comparação dos sistemas OLTP e OLAP acaba por ser bastante frequente, não como tema principal, mas sim aquando a abordagem de apenas um dos temas.

Após abordar estes dois assuntos iniciais, especialmente o OLAP, será mais fácil apresentar o seguinte. O tema que se seguirá irá abordar o conceito e características do ROLAP e do MOLAP.

Apesar de existirem vários tipos de OLAP, este estudo só irá abordar dois deles. A escolha recaiu sobre o ROLAP e o MOLAP por serem os principais tipos de OLAP e como o que levou à escolha do tema foi o facto de profissionalmente me encontrar a trabalhar com dados provenientes destes dois tipos de OLAP foi mais uma razão para os escolher. A decisão de comparar estes dois tipos de OLAP ao nível do *reporting* e, consequentemente, a escolha do software MicroStrategy para, através dele, comparar

relatórios desenvolvidos utilizando fontes de dados ROLAP e MOLAP deveu-se também a questões profissionais, uma vez que é com este software que estou a trabalhar.

Apesar deste tema já ter sido abordado em alguns artigos, não foram muitos os que se focaram apenas na comparação do ROLAP com o MOLAP. Muitos autores ao abordarem outros temas como, por exemplo, sistemas de apoio à decisão, acabaram por se referirem ao ROLAP e MOLAP.

Em 1996, George Colliat comparou o ROLAP com o MOLAP, chegando mesmo a uma conclusão sobre qual o mais adequado (o MOLAP). Em 2000, Hasan, Hyland, Dodds e Veeraraghavan sugerem a comparação entre o ROLAP e o MOLAP como uma questão que acreditavam que seria investigada brevemente. Per Westerlund, na sua tese de Mestrado, em 2008, apresenta uma parte prática, na qual compara a performance do ROLAP e MOLAP chegando à conclusão de que, geralmente, o ROLAP é melhor.

O ano passado, Kazi, Radulovic, Radovanovic, e Kazi (2010) apresentaram um exemplo de implementação de um sistema de DW (*data warehouse*) MOLAP, no qual criaram um cubo OLAP e abordaram o Microsoft Excel como ferramenta de *reporting*. O termo *data warehouse* foi introduzido por Inmon, em 1992, que o definiu como “uma colecção de dados de suporte à gestão das decisões orientados para um assunto, integrados, não voláteis e variantes no tempo” (Kazi et al., 2010), “utilizada principalmente na tomada de decisão organizacional” (Chaudhuri & Dayal, 1997).

O que se pretende ao realizar este trabalho é, em vez de abordar apenas o MOLAP, tal como Kazi, Radulovic, Radovanovic e Kazi (2010) o fizeram, comparar o ROLAP e o MOLAP, tal como George Colliat (1996) e Per Westerlund (2008), mas focando a sua comparação mais ao nível do *reporting*, utilizando para tal a ferramenta de *reporting* MicroStrategy.

A revisão da literatura que se irá apresentar está organizada por temas, começando por uma abordagem sucinta ao OLTP, seguindo-se o OLAP e as diferenças entre os dois e passando para o ROLAP e MOLAP, onde se vai comparando as características de cada um.

2.1. OLTP e OLAP

OLTP é um termo “utilizado para descrever o armazenamento de dados relacionais que é projectado/desenhado e otimizado para actividades transaccionais” (Goff et al., 2009), sendo os sistemas OLTP “o esqueleto dos actuais sistemas de informação, suportando as operações diárias” (Bog et al., 2008). Estes sistemas, que foram desenvolvidos há mais de 30 anos (Bog et al., 2009), automatizam o processamento de dados (Hasan et al., 2000; Chaudhuri & Dayal, 1997) de “tarefas estruturadas e repetitivas” (Chaudhuri & Dayal, 1997).

As aplicações OLTP utilizam na maior parte dos casos os sistemas de gestão de bases de dados relacionais (RDBMS), pois a “normalização das tabelas utilizadas num RDBMS elimina a redundância dos dados e torna mais rápido e simples guardar transacções de negócio” (Hasan et al., 2000).

Com o aumento exponencial da quantidade de dados recolhidos pelas empresas (Asghar & Osama, 2009; Hasan et al., 2000), o OLTP assume um importante papel. Esta importância deveu-se ao facto do OLTP, que se refere a aplicações informáticas que “permitem as operações diárias de negócio de uma empresa” (Bog, Domschke, Mueller, & Zeier, 2009), ter sido criado para “extrair dados de um sistema/base de dados desenvolvido tendo em mente o desempenho da aplicação” (Venkat, n.d.).

Com este aumento da quantidade de dados, também surgiu a necessidade de se criar cada vez mais relatórios que ajudassem os gestores das empresas a analisar a informação rapidamente e a perceberem quais as decisões que deveriam tomar de acordo com os objectivos que eram estabelecidos (John, 2008). No entanto, esta necessidade de criar relatórios fez surgir um problema: as ferramentas de reporting eram muito “pobres no OLTP” (Venkat, n.d.). Por isso, foi criado o OLAP. Este apareceu em 1970, aquando do aparecimento do Express, que foi o primeiro produto que permitia consultas OLAP (Abelló & Romero, n.d.). No entanto, o aparecimento do termo OLAP demorou mais alguns anos a aparecer e só em 1993 foi introduzido por Codd (Cramer, 2006; Abelló & Romero, n.d.). Mas quando Codd introduziu o termo e definiu um conjunto de características que definiriam OLAP, estas foram decididas em conjunto com a Arbor Software. Por isso, recentemente apareceu uma nova definição, FASMI,

(*fast analysis of shared multi-dimensional information*), independente de qualquer tecnologia, dada por Creeth e Pendse (Hasan et al., 2000). No fundo, o termo FASMI é um termo “alternativo e mais descritivo para se referir a OLAP” (Abelló & Romero, n.d.).

Apesar deste termo FASMI, o que irei utilizar será o termo OLAP.

OLAP é um termo que se refere “a aplicações informáticas que permitem efectuar, de forma rápida e partilhada, a análise de informação multidimensional, originária de diversas fontes de dados”, extraíndo conhecimento dos *data marts* que constituem o DW (Abelló & Romero, n.d.).

OLAP “fornece um conhecimento dos dados cujos utilizadores são incapazes de ver apenas com as suas capacidades naturais”, dando “uma visão dos mesmos com diferentes ângulos” (Asghar & Osama, 2009).

Como referido inicialmente, o seu aparecimento foi despoletado principalmente devido ao rápido e exponencial crescimento dos dados na década de 80, o que dificultava cada vez mais a manipulação dos mesmos (Codd, E., Codd, S. & Salley, 1993). No entanto, a importância do OLAP vai além da manipulação de grandes quantidades de dados, pois o OLAP permite aos utilizadores terem “uma visão analítica dos dados, resultando numa tomada de decisão eficaz” (Asghar & Osama, 2009), possibilitando a implementação com sucesso dos SAD (Conn, 2005).

Por isso, refere-se que as aplicações OLAP são SAD interactivos (Pirnaeu & Botezatu, 2010) e cada vez mais utilizados como suporte à decisão (Korobko & Penkova, 2010). As suas principais aplicações são “*reporting* de negócio para vendas, marketing, *reporting* de gestão, gestão de processos de negócio (BPM), orçamento e previsões, *reporting* financeiros e descoberta de conhecimento” (Zhou & Chen, 2009).

A principal característica dos sistemas OLAP é a multidimensionalidade (Abelló & Romero, n.d.), que “implica a capacidade de processar rapidamente os dados de uma base de dados de forma a que as respostas possam ser geradas com rapidez” (TechTarget, 2011a). Na multidimensionalidade as operações são efectuadas sobre um cubo de dados que “fornece aos analistas de dados uma forma intuitiva para navegar através de vários níveis de informação resumida no *data warehouse*” (Lenz & Thalheim, 2009), apresentando como vantagens “esquemas de dados mais compreensíveis para os utilizadores finais e armazenamento específico de técnicas de

acesso que melhoram o desempenho de consultas” (Cramer, 2006). Um cubo de dados pode ser definido como o “conjunto de resultados de consultas agregadas *group-by* em todas as possíveis combinações de dimensões de atributos sobre uma tabela de factos de um *data warehouse*”, podendo também ser definido como uma “representação abstracta de uma projecção da relação dos RDBMS (*Relational database management system*)” (Morfonios & Ioannidis, 2008).

Apesar do termo OLAP estar relacionado com bases de dados multidimensionais, não significam o mesmo, uma vez que as bases de dados multidimensionais são “representadas por um hipercubo ou matriz multidimensional” que pode ser manipulado através de operações de “*slicing e dicing* ao cubo ou de *rolling up e drilling down* às dimensões”, enquanto que o OLAP é o “conjunto de ferramentas que suportam estas manipulações” (Hasan et al., 2000; Cramer, 2006). Nas bases de dados multidimensionais utiliza-se principalmente a linguagem MDX (*Multidimensional Expressions*), tal como nas bases de dados relacionais se utiliza o SQL (Cramer, 2006).

Para além desta vantagem, o OLAP apresenta também uma grande desvantagem. Essa desvantagem prende-se com o facto de não se realizarem análises aos dados em tempo real, efectuando principalmente análises aos dados históricos (Conn, 2005). Análises aos dados de negócio em tempo real são importantes, pois leva a que os processos de negócio sejam optimizados e os SAD melhorados (Conn, 2005). Por isso, há quem defenda que o ideal é combinar o sistema OLAP com o sistema OLTP (Bog et al., 2009; Conn, 2005).

Quem considera que a melhor opção seria juntar estes dois sistemas apresenta como benefícios “a análise dos dados em tempo real”, “*reporting ad-hoc*” e a “optimização das estruturas dos dados e da redundância” que podia ser feito conhecendo antecipadamente as consultas à base de dados (Bog et al., 2009; Conn, 2005).

No entanto, Samuel Conn (2005), baseado na opinião de Helen Thomas e Anindya Datta, defende que não é possível o mesmo modelo lógico e físico da base de dados para OLTP e OLAP, uma vez que os requisitos necessários para OLAP e OLTP são diferentes.

Contudo, também há quem defenda que o ideal é separar os ambientes OLTP e OLAP, uma vez que estes são diferentes. O OLTP trata das necessidades operacionais e

transaccionais e gere dados actuais, automatizando os processos de negócio operacionais, enquanto o OLAP suporta a análise de negócio, servindo de apoio à decisão, e gere dados menos recentes que o OLTP (Conn, 2005; Hasan et al., 2000).

Apesar de “os sistemas OLTP e OLAP serem sistemas separados”, os sistemas OLAP são sistemas dependentes dos sistemas OLTP (Bog et al., 2009) e, por isso, são arquitecturas de bases de dados complementares (Conn, 2005).

OLTP é definido de forma matemática rigorosa, enquanto o OLAP apresenta uma carência de quadro matemático rigoroso (Hasan et al., 2000 e Lenz & Thalheim, 2009). Outras diferenças existentes entre os sistemas OLTP e os sistemas OLAP prendem-se com as operações típicas de cada um (no OLTP utiliza-se a actualização dos dados, enquanto o OLAP é orientado para a análise dos dados) e o nível de detalhe dos dados: os dados provenientes de sistemas OLAP são visualizados de forma agregada, enquanto os sistemas OLTP apresentam dados com maior detalhe e carecem de uma “baixa performance na agregação da informação” (Anzanello, 2002). Ainda relativamente aos dados, nos sistemas OLTP estes são utilizados por todos os colaboradores de uma empresa e podem ser alterados, enquanto nos sistemas OLAP são apenas utilizados pelos gestores das empresas e não são alterados (Anzanello, 2002).

Após se ter abordado algumas das diferenças entre OLTP e OLAP, irá apresentar-se de seguida mais algumas.

Quadro 2.1. *Diferenças entre OLTP e OLAP*

	OLTP	OLAP
Uso/Utilização	Aplicação específica	Suporte à decisão
Workload	Predefinido	Imprevisível
Acesso	Leitura/Escrita	Apenas Leitura
Estrutura da consulta	Simples	Complexa
Registos por operação	Dezenas/Centenas	Milhares/Milhões
Número de utilizadores	Milhares/Milhões	Dezenas/Centenas

Fonte: Abelló & Romero, n.d.

2.2. Importância OLAP

A utilização dos SAD são uma forma encontrada pelas empresas de aumentarem a produtividade e obterem vantagem competitiva, tornando os seus “utilizadores de negócio mais auto-suficientes e posicionando a organização para responder mais rapidamente à exigência de um mercado” (Hart & Porter, 2004) onde há cada vez mais concorrência (Hart & Porter, 2004; Morfonios & Ioannidis, 2010). Os sistemas OLAP, ao ajudarem na tomada de decisão passam a ser elementos de suporte à decisão e, por consequência, também eles cada vez mais importantes (Hasan et al., 2007; Bhan et al., 2009; Cuzzocrea, 2011; Dehne et al., 2003; Dehne et al., 2005), permitindo aos gestores tomarem as suas decisões com base em “informações de forma rápida e precisa, com alta flexibilidade e desempenho” (Cramer, 2006; Asghar & Osama, 2009). Estas informações irão permitir aos gestores obterem “*insights* dos dados através de um acesso rápido, consistente, interactivo a uma grande variedade de possíveis pontos de vista da informação” (Lehn et al., 1997), tendo “capacidade de analisar vários cenários, fazer análises *what if*, consultas *ad hoc*, previsões e descobrir tendências” (Cramer, 2006).

OLAP permitiu a criação de aplicações de negócio essenciais para as empresas, como por exemplo, “análise, planeamento, orçamento e avaliação do desempenho das vendas e do marketing” (Dehne et al., 2006).

Em 1997, McCulley afirmou que, actualmente, a implementação de maior sucesso de sistemas de apoio à decisão é OLAP, “permitindo uma rápida análise dos dados e *reporting* de grandes fontes de dados” (McCulley, 1997).

2.3. Tipos de OLAP

Os sistemas OLAP podem ter várias arquitecturas, sendo as principais o ROLAP e MOLAP.

O ROLAP é uma arquitectura em que o armazenamento de dados multidimensionais é feito em bases de dados relacionais (Pirnau & Botezatu, 2010), enquanto no MOLAP o armazenamento é feito em bases de dados multidimensionais.

O ROLAP depende de tecnologia de bases de dados relacionais, tendo de utilizar linguagem SQL (*Structured Query Language*) para implementar operações que utilizem estruturas de dados multidimensionais, no entanto esta linguagem não é apropriada para operações multidimensionais (o SQL não é suficientemente poderoso ou flexível para suportar as capacidades OLAP) e dificulta a sua optimização (Pirnau & Botezatu, 2010; Buzydlowski, Song, & Hassel, 1998; Abelló & Romero, n.d.). Isto leva a que os sistemas ROLAP apresentem uma baixa performance, nomeadamente muito lenta, comparativamente com os sistemas MOLAP que, por sua vez, apresentam uma boa performance e ainda permitem cálculos complexos (Conn, 2005; Abelló & Romero, n.d.).

As bases de dados relacionais são compostas por tabelas de factos e uma tabela por dimensão (Figura 2.1.) e utiliza-se como modelo lógico um *star schema* desnormalizado para atingir uma performance aceitável (Colliat, 1996). Por tabelas de factos entende-se “tabelas da base de dados que contêm dados numéricos e que podem ser agregados ao longo de uma ou mais dimensões” (MicroStrategy, 2009b), enquanto as tabelas de dimensões são tabelas onde se encontram “armazenados os atributos ou dimensões” que descrevem os factos da tabela de factos, sendo uma dimensão uma “coleção de informação de referência sobre um evento que pode ser medido, i.e., factos” (TechTarget, 2011b). Relativamente ao modelo lógico, este pode ser *star schema* ou *snowflake*. *Star schema* é quando “cada tabela de factos está associada a dimensões, assemelhando-se o diagrama a uma estrela” (TechTarget, 2011d) e o *snowflake* quando “tal como no *star schema*, cada tabela de factos está associada a dimensões, mas depois essas dimensões estão relacionadas a outras dimensões, ramificando-se num padrão floco de neve” (TechTarget, 2011c).

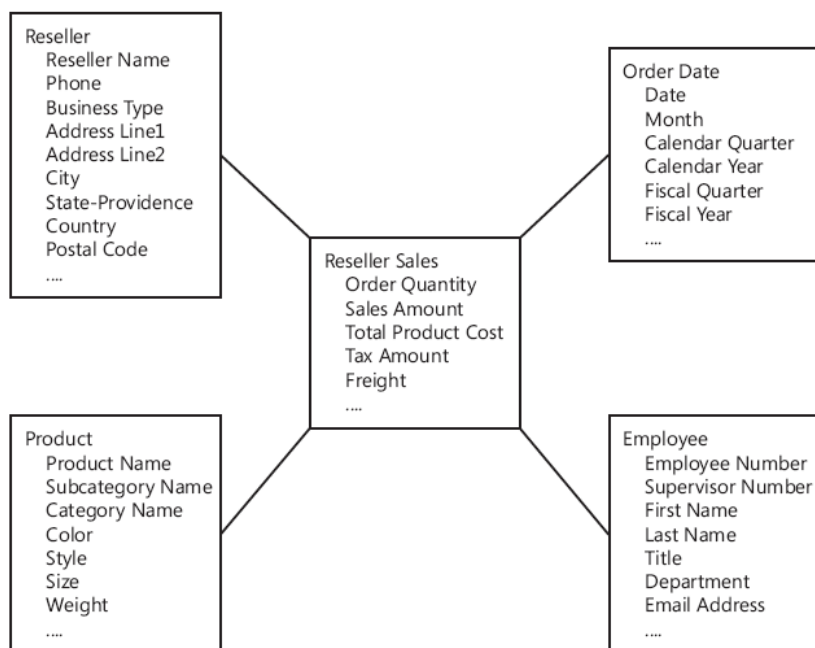


Figura 2.1. Exemplo de uma tabela de factos e quatro tabelas de dimensões, tendo como modelo lógico o *star schema*.

Fonte: Clay, Consulting & Smith, 1999

As bases de dados multidimensionais designam-se por cubos (Figura 2.2.), os quais permitem uma análise rápida dos dados (Kazi et al., 2010; Abelló & Romero, n.d.). “Num cubo cada valor dos dados ou do facto ocupa uma célula indexada por um único conjunto de valores de dimensão” (Hasan et al., 2000). Para um modelo multidimensional deve-se também utilizar como modelo lógico o *star schema* (Kazi et al., 2010; Buzydlowski et al., 1998).

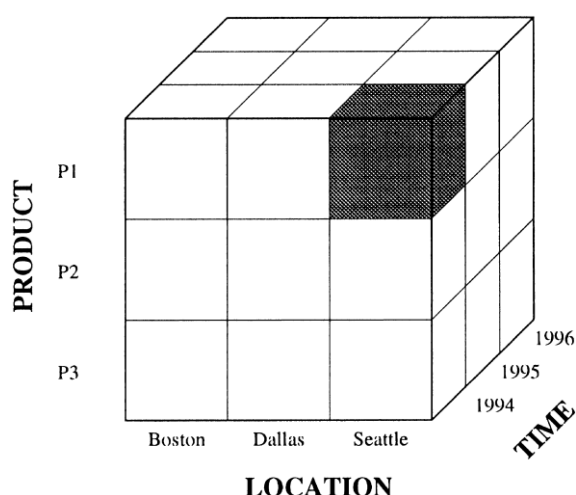


Figura 2.2. Exemplo de um cubo.

Fonte: Datta & Thomas, 1999

O ROLAP apresenta uma boa standardização das tecnologias de bases de dados (Buzydlowski et al., 1998; Abelló & Romero, n.d.; Hasan et al., 2000), já no MOLAP não existe uma standardização (Buzydlowski et al., 1998).

Como o ROLAP utiliza consultas relacionais, permite que os dados permaneçam na terceira forma normal tal como acontece num sistema OLTP (Conn, 2005). Também “accede directamente aos dados nas bases de dados relacionais do DW e consegue trabalhar com grandes quantidades de dados pois não transfere os dados para o DW” (Kazi et al., 2010; Abelló & Romero, n.d.).

Em comparação, o MOLAP é limitado na quantidade de dados que consegue guardar/manipular (Abelló & Romero, n.d.), apresenta problemas de escalabilidade, não conseguindo trabalhar com grandes quantidades de dados (Buzydlowski et al., 1998).

Como o ROLAP accede directamente à base de dados relacional, torna-se fácil adicionar novas dimensões, para além do facto dos dados estarem sempre disponíveis. (Kazi et al., 2010).

No caso do MOLAP realizar quaisquer alterações, tais como adicionar dimensões, é muito complicado, pois, devido à falta de flexibilidade, requer a reconstrução do mesmo (Kazi et al., 2010; Korobko & Penkova, 2010; Conn, 2005).

No entanto, aceder aos dados directamente também tem inconvenientes, como quando se pretende fazer operações de *drill down* nos dados, para que isto aconteça em tempo real é preciso possuir-se muito poder de processamento e por isso hardware caro

(Hasan et al., 2000). Já os sistemas MOLAP “estão focados na flexibilidade, na optimização de técnicas de armazenamento e no conceito de transacção” (Pirnaeu & Botezatu, 2010) e por isso pode ser utilizado “para acelerar o tempo de performance de uma consulta OLAP, uma vez que os dados estão construídos num persistente cubo dimensional que é separado dos dados transaccionais” (Conn, 2005).

Outras diferenças entre MOLAP e ROLAP são que o MOLAP armazena uma cópia dos dados carregados (agora é uma vantagem) fazendo com que os dados possam ser limpos e várias agregações realizadas durante o carregamento dos mesmos para um maior desempenho e flexibilidade aquando o utilizador final aceder aos dados multidimensionais (Hasan et al., 2000).

Ao contrário dos sistemas ROLAP, os sistemas MOLAP não permitem pedidos *ad hoc* eficazmente, uma vez que eles são optimizados para operações multidimensionais (Pirnaeu & Botezatu, 2010).

Tanto os sistemas ROLAP como MOLAP são ideais para “aplicações com pouca volatilidade dos dados, mas para aplicações com elevada volatilidade dos dados apenas o sistema ROLAP é ideal” (Pirnaeu & Botezatu, 2010).

Normalmente, as vantagens do MOLAP são as desvantagens do ROLAP e vice-versa (Buzydlowski et al., 1998).

Pode-se dizer que os sistemas ROLAP são adequados para guardar grandes quantidades de dados, usando um processamento paralelo e tecnologias separadas, enquanto os sistemas MOLAP são adequados para aplicações departamentais com pequenos volumes de dados (Pirnaeu & Botezatu, 2010; Kazi et al., 2010).

No entanto, Colliat (1996) que escreveu um artigo no qual compara o ROLAP com o MOLAP, no final desse artigo toma uma posição mais drástica e afirma que o ROLAP não é apropriado para o OLAP “devido ao elevado número de I/O necessários para realizar um simples *drill down* e computação” (Colliat, 1996). Este autor sugere ainda a utilização do MOLAP (Colliat, 1996), pois chegou à conclusão que este ocupava menos espaço do que o ROLAP, devolvia os dados “8 a 200 vezes mais rápido” e calculava os dados derivados também mais rápido (Colliat, 1996).

Também Werterlund, na sua tese de Mestrado, onde compara e testa o ROLAP com o MOLAP acaba por concluir que “o MOLAP apresenta uma melhor performance”, mas “se os ganhos ao nível da performance usando o MOLAP não forem

muito grandes, então não existe nenhuma razão para escolher o MOLAP” (Westerlund, 2008).

Estas duas arquitecturas estão a ser apresentadas neste capítulo como sendo arquitecturas concorrentes, até porque é pressuposto que irá estar presente ao longo do restante trabalho, mas também podem ser vistas como arquitecturas de bases de dados complementares (Hasan et al., 2000).

Existem ainda autores que sugerem a implementação de outra arquitectura que não o ROLAP nem o MOLAP. O Buzydlowski, Song e Hassel (1998) sugerem a implementação de O3LAP (*object-oriented online analytic processing*) utilizando conceitos OO (*object-oriented*), pois consegue maximizar as vantagens e minimizar as desvantagens do MOLAP e ROLAP. Conn (2005) sugere a implementação do DOLAP (*dynamic online analytical processing*).

Apesar destas arquitecturas sugeridas, no trabalho futuro só se irá abordar as arquitecturas ROLAP e MOLAP.

3. REPORTING

3.1. OLAP Reporting

Como o objectivo deste estudo é estabelecer a comparação entre dois tipos de sistemas OLAP para o propósito de *reporting* era importante abordar as actividades de reporting e a sua classificação.

Antes de abordar este assunto e com o objectivo de permitir uma maior compreensão do mesmo, irá realizar-se a distinção entre actividades operacionais e actividades baseadas em análise, orientadas à decisão.

Quadro 3.1. *Diferenças entre Actividades Operacionais e Actividades Baseadas em Análise, Orientadas à Decisão*

Actividades Operacionais	Actividades baseadas em Análise, orientadas à Decisão
Mais frequentes	Menos frequentes
Mais previsíveis	Menos previsíveis
Menores quantidades de dados acedidos para consulta	Maiores quantidades de dados acedidos por consulta
Consulta principalmente de dados primitivos	Consulta principalmente de dados derivados
Exige principalmente dados actuais	Exige dados passados, presentes e projectados
Pouca ou nenhuma derivação complexa	Muitas derivações complexas

Fonte: Cramer, 2006

Após esta distinção, irá então apresentar-se algumas classificações das actividades de *reporting*. Em 2000, Inmon distingue os sistemas OLAP em dois tipos de *reporting*: operacionais (inclui as mais recentes informações dentro do sistema OLTP, que permitem analisar e suportar as actividades do dia-a-dia em detalhe) e informacionais (suportam a tomada de decisão estratégica e a longo prazo, sendo a enorme quantidade de dados resumida) (Bog et al., 2008; Bog et al., 2011; SourceMedia, 2011).

Em 2006, White, em vez de dois, distingue três tipos de relatórios: estratégicos, táticos e operacionais. Tal como acontecia com os relatórios informacionais definidos por Inmon, os relatórios estratégicos e táticos são ambos baseados em informações históricas. A diferença entre os relatórios estratégicos e os táticos encontra-se no período de tempo analisado: os primeiros analisam meses a anos e os segundos dias da semana a meses. Os relatórios operacionais, tal como Inmon o fez, são definidos também como os relatórios para otimizar operações diárias (Bog et al., 2011).

Também em 2006, Frolick e Ariyachandra defendem que os dados em tempo útil suportam tanto a tomada de decisão operacional como a estratégica e tática (Bog et al., 2011).

3.2. Possíveis Abordagens para Relatórios

Tal como abordado anteriormente existem vários tipos de relatórios, não existindo um total consenso sobre a sua classificação.

Cada relatório depende dos objectivos que se pretendem analisar. Nos relatórios a desenvolver numa futura fase deste estudo irá optar-se, segundo a classificação de White, por criar relatórios estratégicos.

Nesses futuros relatórios será interessante tentar reproduzir alguns dos seguintes relatórios:

1. Relatórios simples como, por exemplo, uma evolução do número de produtos vendidos por mês com o intuito de perceber se o número de produtos vendidos está a aumentar ou diminuir ou se há meses em que há um maior ou menor número.
2. Relatórios mais complexos, por exemplo uma evolução do número de produtos vendidos por mês, por tipo de produto e por loja para analisar quais os tipos de produtos mais vendidos em cada loja e se estes se mantêm ao longo dos meses.

3. Relatórios para comparar os valores actuais com os homólogos e conhecer a variação entre estes de forma a, por exemplo, comparar o número de produtos vendidos num determinado mês com o mesmo mês do ano anterior.

4. Relatórios onde os dados estão filtrados, por exemplo, visualiza-se só os dados dos meses do último ano ou só as lojas que estão localizadas em determinada região geográfica.

5. Relatórios onde se visualizam todos os dados, por exemplo, os dados dos meses dos vários anos, todos os tipos de produtos ou de todas as lojas.

6. Relatórios onde se escolhe na altura quais os dados a visualizar, por exemplo qual(ais) os meses ou quais os tipos de produtos que se pretende visualizar.

7. Relatórios onde se escolhe visualizar, por exemplo, as lojas ou os produtos que têm mais produtos vendidos.

Vai ser com base nestas possíveis análises que se irá criar futuramente os relatórios.

4. METODOLOGIA

4.1. Investigação-Acção

“Escolher a melhor abordagem para um estudo depende do objectivo do mesmo e do encaminhamento das questões de investigação” (Tabatabaei, 2009). Por isso, neste estudo irá adoptar-se uma abordagem qualitativa, à semelhança da metodologia adoptada em estudos análogos (Conn, 2005; Hasan et al., 2000). A epistemologia adoptada será a interpretativa e mesmo existindo vários métodos qualitativos, neste caso irá ser utilizado o método de investigação-acção.

O método de investigação-acção é caracterizado por ser um ciclo repetitivo cujo objectivo é “produzir soluções tanto para o problema actual como para o avanço do conhecimento sobre o domínio do problema em investigação, produzindo resultados de pesquisa relevantes” (Gregg, 2003), “pois baseia-se em acções concretas, destinadas a resolver uma situação problema de imediato, enquanto cuidadosamente informa a teoria” (Baskerville, 1999).

A escolha recaiu sobre este método por ser “uma forma de investigação centrada no problema que faz a ligação entre a teoria e a prática” (Sanford, 1970), que é o que se pretende neste trabalho, onde se irá realizar uma aplicação prática dos temas inicialmente abordados, mas também por este método ser cada vez mais utilizado na área de sistemas de informação (Conn, 2005; Hasan et al., 2000; Myers, 2010; Baskerville, 1999).

A investigação acção tem as seguintes etapas: “diagnóstico do problema, planeamento de acções, tomada das acção, avaliar os resultados e especificar a aprendizagem” (Gregg, 2003).

4.2. Apresentação da Empresa

A empresa à qual se recorreu, solicitando a utilização dos dados para a realização deste estudo, é uma empresa portuguesa criada há mais de 20 anos, a qual se designará por Gestão de Recursos.

Esta empresa está vocacionada para a área de Sistemas e Tecnologias de Informação tendo, para além da componente técnica, um elevado conhecimento do negócio em causa.

4.3. Apresentação/Caracterização dos Dados

Os dados que se irão utilizar foram disponibilizados pela empresa Gestão de Recursos. Correspondem à área de Marketing, sendo a base de dados de um programa de fidelização de clientes e, por isso, apresenta dados referentes aos seus clientes e actividades efectuadas pelos mesmos. Por conter os dados dos clientes, os nomes destes e outras informações que os identifiquem não poderão ser apresentados.

Estes dados encontram-se sobre uma base de dados relacional, utilizando o software Microsoft SQL Server 2008 R2, nomeadamente o SQL Server Management Studio. Esta base de dados tem cerca de 50 tabelas (das quais 10 são tabelas de factos e as restantes tabelas de dimensão) que guardam 50 milhões de registos. Destas tabelas só 12 (3 tabelas de factos e 9 tabelas de dimensões) serão utilizadas neste estudo para a criação de relatórios. O modelo lógico desta base de dados é o *snowflake*.

Esta base de dados foi utilizada para criar um cubo OLAP, utilizando para tal o software SSAS (SQL Server Analysis Services) do Microsoft SQL Server 2008 R2. Este cubo armazena mais de 100 métricas e cerca de 60 dimensões. Destas, vão ser utilizadas 13 métricas e 9 dimensões.

As duas bases de dados descritas anteriormente irão ser utilizadas para a criação de relatórios, através do software MicroStrategy versão 9.0.2, nomeadamente do MicroStrategy Desktop. Este software permite utilizar duas arquitecturas: relacional OLAP e multidimensional OLAP, apresentando-se de seguida um esquema de cada uma delas com as fases e respectivas funcionalidades até à fase de *reporting* (*Presentation Layer*).

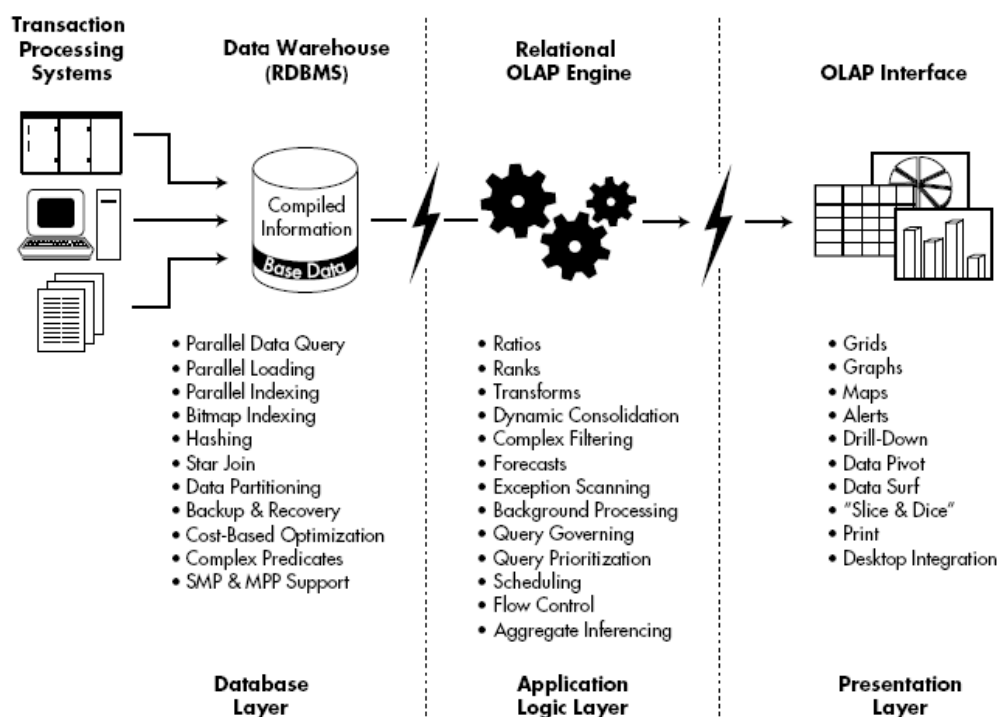


Figura 4.1. Arquitetura ROLAP.

Fonte: MicroStrategy, 1995

A Figura 4.1. descreve a arquitetura ROLAP que “accede aos dados armazenados no *data warehouse* para fornecer análises OLAP” (MicroStrategy, 1995). Nesta arquitectura deve-se começar por definir o modelo de dados do *data warehouse* para depois carregar os dados de diferentes sistemas de processamento numa base de dados. Se for necessário para o modelo de dados, a base de dados é executada para que os dados sejam agregados. De seguida serão criados os índices que irão “optimizar o tempo de acesso às consultas à base de dados” (MicroStrategy, 1995), passando “os utilizadores finais a poderem submeter análises multidimensionais ao ROLAP” (MicroStrategy, 1995). Estas análises multidimensionais serão transformadas

dinamicamente em execuções SQL, sendo o “SQL submetido à base de dados relacional para processamento e retornando os resultados da consulta relacional que, por sua vez, serão *cross-tabulated* para ser retornado o resultado multidimensional ao utilizador final” (MicroStrategy, 1995).

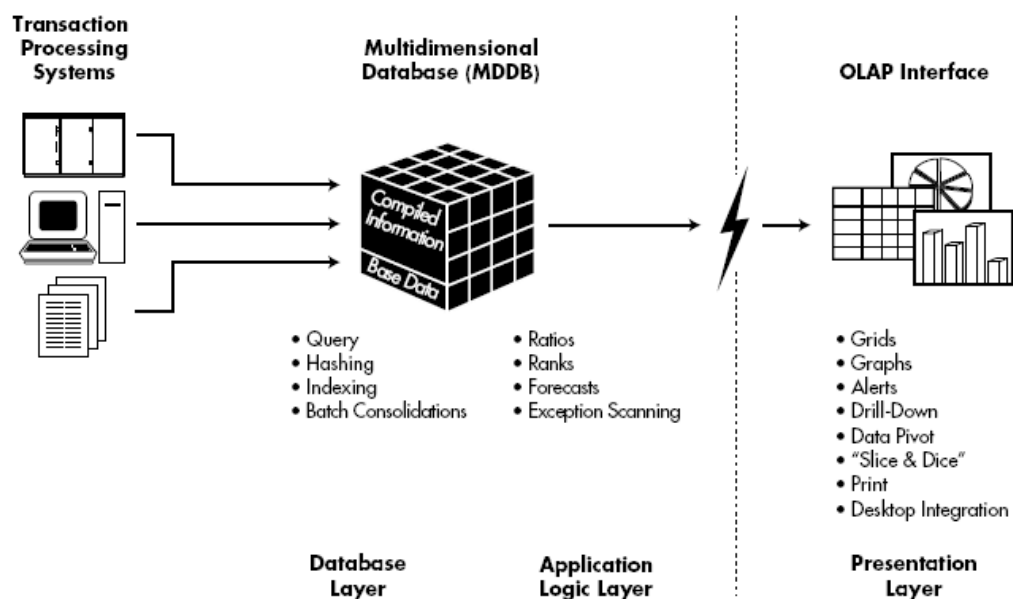


Figura 4.2. Arquitetura MOLAP.

Fonte: MicroStrategy, 1995

Na Figura 4.2. é apresentada a arquitetura MOLAP. Nesta arquitetura, a informação carregada nas bases de dados multidimensionais é de diferentes sistemas operacionais. Após o carregamento realizam-se vários cálculos para “agregar as dimensões ortogonais e preencher as estruturas *array* MDDB” (MicroStrategy, 1995). De seguida, a base de dados multidimensional já pode ser utilizada, podendo os utilizadores finais fazer pedidos OLAP aos quais a *application logic layer* retornará os dados armazenados. Nesta arquitetura a *database layer* (que armazena os dados, acede e retorna os processos) e a *application logic layer* (executa todos os pedidos OLAP) encontram-se como fazendo parte da base de dados multidimensional.

Os relatórios a construir em cada uma das duas arquiteturas irão ter diferentes funcionalidades e níveis de detalhe. Para cada funcionalidade e nível de detalhe existirão dois relatórios idênticos, tendo apenas como diferença a fonte de dados: um

deles terá como fonte de dados o relacional e o outro o multidimensional. Estes relatórios serão comparados nos seguintes aspectos: o tempo que demoram até retornar os dados e o espaço que ocupam. Ao comparar estes relatórios espera-se descobrir se existem diferenças entre eles e, se tal se verificar, quais são as principais.

4.4. Descrição das Etapas na Criação dos Relatórios

Após a apresentação dos dados e dos softwares utilizados irá passar-se à descrição das tarefas necessárias até à criação dos relatórios, apresentando-se na Figura 4.3. um esquema com essas etapas para o ROLAP e MOLAP.

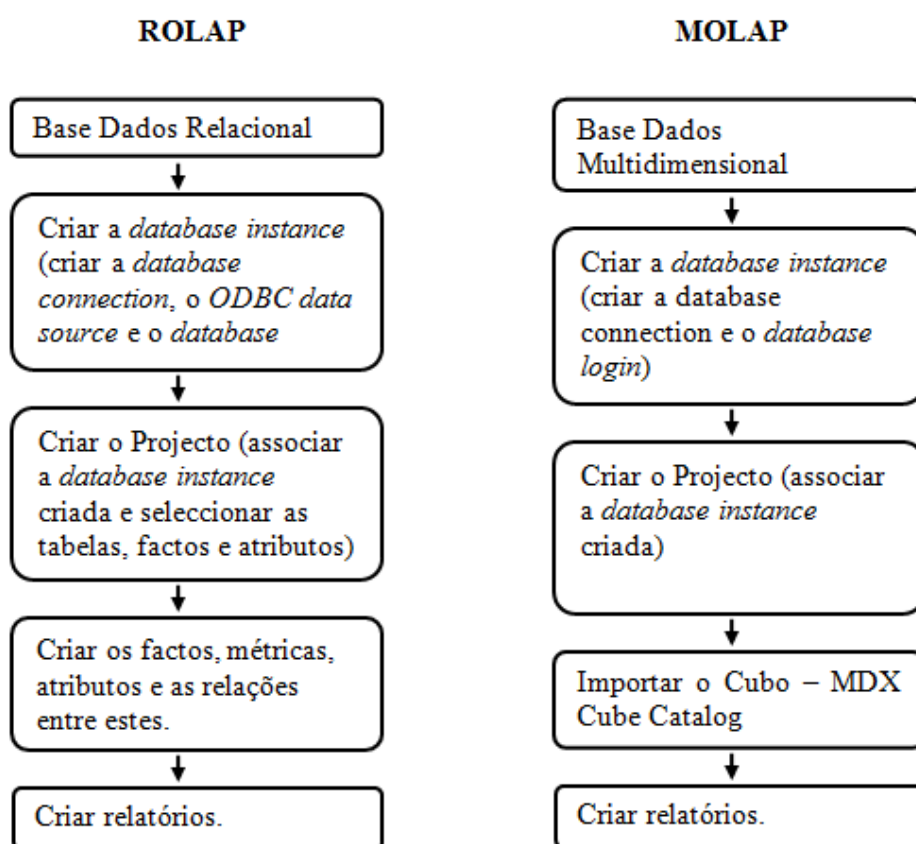


Figura 4.3. Esquema das etapas na criação dos relatórios.

4.4.1. Base de dados.

Uma vez que tanto a base de dados relacional como a multidimensional já estavam criadas e disponíveis para serem utilizadas no desenvolvimento de actividades de *reporting*, não foi necessário tomar qualquer acção no Microsoft SQL Server 2008 R2. No entanto, o mesmo não aconteceu no software MicroStrategy.

No MicroStrategy já existia um projecto, designado Marketing, no qual já existiam alguns relatórios desenvolvidos sobre o cubo OLAP que se iria utilizar. Por isso, utilizou-se o projecto existente para desenvolver os relatórios que se pretendia, não tendo sido necessário realizar todos os passos descritos na Figura 4.3.

4.4.2. Database instance.

De seguida teve de se criar/confirmar a ligação do MicroStrategy às bases de dados. Para tal é necessário que existam duas *databases instances*, uma que liga às tabelas e outra ao cubo OLAP. Por *database instance* entende-se a ligação do MicroStrategy à base de dados e, aquando da sua criação, é preciso definir o *DSN data warehouse*, *Login ID* e *password* (MicroStrategy, 2009b).

A *database instance* que faz a ligação ao cubo OLAP, que tem de ser do tipo Microsoft Analysis Services 2008, já se encontrava criada, por isso não foi necessário tomar nenhuma acção. No entanto, a *database instance* que faz a ligação às tabelas ainda não tinha sido criada, por isso foi necessário criar um *ODBC data source* para se conseguir criar um DSN e então utilizá-lo na criação da *database instance*, que é do tipo Microsoft SQL Server 2008. Por ODBC (*Open Database Connectivity*) entende-se “*open standard* com o qual os computadores clientes podem comunicar com os servidores das bases de dados relacionais” (MicroStrategy, 2009b).

4.4.3. Projecto.

Após a criação das *databases instances*, no caso dos dados sobre o relacional foi necessário adicionar, ao projecto de Marketing no MicroStrategy, as tabelas que se iriam utilizar, de forma a que se pudesse criar factos, métricas, atributos e as relações entre eles. Importa ainda referir a definição de cada um deles. Factos são “objectos que representam a coluna de uma tabela do *data warehouse* e contêm números básicos ou agregados” (MicroStrategy, 2009b), enquanto as métricas são “um cálculo de negócio definido por uma expressão que é construída com funções, factos, atributos ou outras métricas” (MicroStrategy, 2009b). Atributos são “um nível de dados definido pelo arquitecto do sistema e associado a uma ou mais colunas de uma tabela *lookup* do *data warehouse*” (MicroStrategy, 2009b) e uma tabela *lookup* é nada mais, nada menos do que uma tabela da base de dados com a descrição das dimensões. Quando se refere criar ligações entre os atributos é estabelecer uma relação de pai-filho entre os mesmos (MicroStrategy, 2009b). Estas ligações foram criadas de acordo com as que se encontravam estabelecidas no cubo OLAP, as quais foram visualizadas através do SQL Server Business Intelligence Development Studio. Depois de se ter criado os factos, métricas, atributos e relações entre estes começou a criar-se os relatórios sobre o relacional.

No caso dos dados sobre o cubo OLAP a situação foi diferente. Como já se referiu anteriormente, a *database instance* já estava criada e o cubo também estava disponível para ser utilizado na criação de relatórios no MicroStrategy.

4.4.4. Relatórios.

Depois de estarem criados os relatórios de ambas as fontes de dados, estes irão ser corridos mais do que uma vez (mais precisamente cinco vezes) e em dias diferentes,

de forma a evitar que se, durante o tempo que estão a correr, alguma vez forem afectados por um factor externo se consiga detectar através da comparação dos resultados do mesmo relatório corrido em dias diferentes. Importa ainda referir que sempre que algum dos relatórios é executado irá apagar-se, de seguida, a *cache* para que esta não venha a influenciar os resultados. Esta decisão de apagar a *cache* foi tomada para permitir que quando se volte a executar o relatório, este volte a correr em vez de retornar o que estava armazenado em *cache*. Por *cache* entende-se “a conservação de um especial armazenamento de dados recentemente acedidos de forma a que um futuro acesso seja mais rápido” (MicroStrategy, 2009b).

Ao executar os vários relatórios pretende-se registar o tempo que demoram a retornar os dados e o espaço que ocupam. Para se obter o tempo de execução de cada um, depois de cada execução vai-se ao SQL/MDX View e regista-se o tempo. Para se conhecer o espaço que cada um ocupa, depois de cada execução vai-se ao menu Administration e dentro de System Monitors irá aparecer a opção Caches, onde se visualizava o tamanho de cada relatório.

Importa ainda referir que cada relatório irá retornar os mesmos dados tanto quando se utiliza como fonte de dados as tabelas ou o cubo.

4.5. Apresentação dos Relatórios

Como anteriormente se explicou os procedimentos necessários para a criação dos relatórios, agora irão apresentar-se os mesmos.

Ao todo são 18 relatórios compostos por diferentes atributos e métricas. Este número de relatórios teve apenas a ver com as diversas funcionalidades que permitiam a sua criação no ROLAP e no MOLAP.

4.5.1. Relatórios 1 a 5.

Os primeiros cinco relatórios têm apenas atributos e métricas (tendo métricas compostas, condicionais e transformadas). Entende-se por métricas compostas as métricas “criadas para combinar uma ou mais métricas com um ou mais operadores matemáticos ou constantes” (MicroStrategy, 2009b). As métricas condicionais são métricas “que contêm o seu próprio filtro que é completamente separado e independente de qualquer critério de filtragem especificado num *report filter*” (MicroStrategy, 2009a). Por fim, as métricas de transformação são “métricas que têm aplicadas as propriedades de uma transformação” (MicroStrategy, 2009b), entendendo-se por transformação um “objecto que encapsula uma regra de negócio utilizado para comparar resultados de diferentes períodos de tempo” (MicroStrategy, 2009b), como é o caso do período de tempo referente ao Mês Homólogo.

O que torna estes cinco relatórios diferentes entre eles é mesmo o número de métricas e atributos que cada um tem. Com eles pretende-se ver se, de um modo geral, irá haver diferenças entre os relatórios das tabelas e do cubo.

Quando se refere os tipos de métricas presentes nos relatórios, esta classificação é feita com base nas métricas criadas no relacional, por exemplo, nos relatórios, quando se tem uma métrica transformada (como é o caso do Homólogo), esta só é uma métrica transformada no relacional, uma vez que no cubo será uma métrica filtrada para um elemento de um atributo (neste caso o atributo Calculated Date).

4.5.1.1. Relatório 1.

Este relatório tem apenas um atributo e uma métrica, a qual se designará por métrica A.

O atributo, que iremos chamar de GP, tem cerca de 120 elementos, todos eles com valor quando cruzados com a métrica A.

A métrica é uma métrica composta: é a soma de duas métricas condicionais. Estas duas métricas condicionais têm factos que utilizam ambos a função SUM sobre

uma coluna de uma das tabelas de factos. Estas duas métricas condicionais são-no porque têm um filtro a excluir um elemento de um atributo que designaremos por TS e outro elemento do atributo IL.

4.5.1.2. Relatório 2.

Este relatório é igual ao relatório anterior, tendo apenas mais um atributo, o atributo Month.

Este atributo Month tem 49 elementos, i.e., apresenta dados de Janeiro 2007 a Janeiro 2011. Também não pertence à mesma dimensão do atributo GP, ou seja, não existe nenhuma relação entre estes dois atributos. Todas as combinações do atributo Month e GP retornam valores para a métrica A.

4.5.1.3. Relatório 3.

O relatório 3 tem apenas um atributo (Month) e 15 métricas, sendo que destas 15, 5 são métricas que apresentam o valor actual para cada mês, outras 5 calculam o valor para o mês Homólogo e as restantes 5 a Variação do valor do mês actual em relação ao mês Homólogo.

Uma das métricas, a que se chamará de B, é uma métrica condicional que tem aplicado um filtro para excluir um elemento de um atributo que designaremos por TS e outro elemento do atributo IL. O facto utilizado por esta métrica é um SUM sobre uma coluna de uma das tabelas de factos.

Outra das métricas presentes neste relatório, a qual se designará por C, é uma métrica composta, sendo a divisão entre a métrica B e uma outra métrica composta que,

por sua vez, é a soma entre duas métricas condicionais às quais se aplica um filtro para excluir um elemento do atributo TS e outro elemento do atributo IL. Estas quatro métricas condicionais utilizam todas factos com a função SUM sobre uma coluna da tabela de factos. Esta métrica é uma percentagem, por isso apresenta os valores em percentagem (%).

Uma outra métrica, a que se chamará de D, é uma métrica composta por duas métricas: uma delas é uma métrica condicional e utiliza um facto cuja função é o SUM sobre uma coluna da tabela de factos; e a outra é uma métrica composta por duas métricas condicionais, em que ambas possuem factos cuja função é o SUM. A estas métricas condicionais é aplicado um filtro para excluir um elemento do atributo TS e outro elemento do atributo IL.

A métrica E, que é composta por outras duas condicionais que, por sua vez, utilizam factos com a função COUNT sobre uma coluna da tabela de factos. A ambas é aplicado um filtro para excluir um elemento do atributo TS e outro elemento do atributo IL.

Outra métrica utilizada é a métrica A, já descrita anteriormente.

As métricas com o Homólogo, no caso dos dados do cubo, são métricas criadas no cubo, em que se filtra a métrica para um elemento do atributo Calculated Date (neste caso para o elemento Homólogo). Este atributo possui os elementos Homólogo e Variação, o que permite que quando cruzado com uma métrica retorne os valores dos Homólogos e das Variação.

Nos relatórios sobre as tabelas, os Homólogos são criados com base em transformações e as Variações são métricas compostas com a seguinte fórmula:

$$\frac{\text{Métrica Actual} - \text{Métrica Homólogo}}{\text{Métrica Homólogo}}$$

Sendo que esta fórmula é aplicada a cada uma das 5 métricas presentes neste relatório (métrica A, B, C, D e E), como por exemplo,

$$\frac{\text{Métrica A} - \text{Métrica Homólogo A}}{\text{Métrica Homólogo A}}$$

Todas as métricas apresentam valores para os vários meses, tendo, no entanto, em conta que para as métricas do Homólogo e Variação só existem valores a partir do ano de 2008.

Month	Metric A	Metric B	Metric C	Metric D	Metric E	Hom Metric A	Hom Metric B	Hom Metric C	Hom Metric D	Hom Metric E	Var Metric A	Var Metric B	Var Metric C	Var Metric D	Var Metric E
08JAN	90,173	660,022	108.3%	647,812	1,089	87,883	608,346	103.1%	605,117	638	50002.9%	50047.7%	5.0%	50039.4%	50326.8%
08FEB	77,101	648,211	67.3%	720,133	1,033	68,487	598,947	272.7%	536,291	691	50012.7%	50049.8%	(75.3%)	50165.8%	50179.1%
08MAR	84,240	656,676	326.2%	548,027	1,162	69,968	610,035	327.9%	533,557	726	50020.5%	50042.4%	(0.5%)	50013.2%	50192.9%
08APR	98,396	657,231	209.8%	574,940	1,035	71,433	593,882	241.8%	538,827	770	50038.0%	50067.5%	(13.2%)	50038.5%	50098.1%
08MAY	96,374	659,114	89.8%	677,089	1,022	72,567	602,261	246.1%	541,555	792	50033.0%	50055.6%	(63.5%)	50132.5%	50078.8%
08JUN	103,376	630,677	228.9%	557,091	1,085	72,957	607,829	282.8%	538,135	768	50042.0%	50021.2%	(19.0%)	50017.6%	50118.3%
08JUL	67,248	649,261	261.0%	557,178	1,156	68,679	639,557	256.7%	554,373	858	49998.4%	50007.0%	1.7%	50002.0%	50083.2%
08AUG	89,796	625,758	210.9%	559,615	1,033	82,163	628,732	267.9%	548,046	765	50009.3%	49997.7%	(21.3%)	50009.0%	50101.1%
08SEP	86,235	659,322	271.4%	558,695	1,209	77,234	674,918	373.8%	546,793	861	50011.7%	49991.1%	(27.4%)	50006.8%	50096.4%
08OCT	89,790	677,675	343.8%	551,679	1,115	76,278	695,783	454.6%	543,064	959	50017.8%	49990.8%	(24.4%)	50004.4%	50034.0%
08NOV	88,420	650,495	315.2%	547,744	1,033	77,989	734,305	363.6%	564,446	919	50013.5%	49964.2%	(13.3%)	49992.9%	50027.2%
08DEC	91,341	612,324	232.8%	548,255	1,182	78,131	596,431	142.3%	567,777	889	50017.0%	50016.5%	63.6%	49979.8%	50075.3%
09JAN	96,678	652,702	48.4%	815,723	132,727	90,173	660,022	108.3%	647,812	1,089	50007.3%	49995.4%	(55.3%)	50104.9%	72349.4%
09FEB	88,638	665,805	52.9%	813,433	118,320	77,101	648,211	67.3%	720,133	1,033	50015.1%	50011.9%	(21.4%)	50063.0%	72005.1%
09MAR	101,999	777,714	31.6%	1,378,137	145,784	84,240	656,676	326.2%	548,027	1,162	50021.2%	50077.3%	(90.3%)	50529.8%	71846.2%
09APR	108,700	777,155	26.1%	1,561,965	169,968	98,396	657,231	209.8%	574,940	1,035	50010.5%	50076.3%	(87.6%)	50627.8%	81576.3%
09MAY	90,480	719,573	70.2%	812,927	156,727	96,374	659,114	89.8%	677,089	1,022	49993.9%	50038.0%	(21.9%)	50085.4%	79828.5%
09JUN	92,015	692,848	60.8%	817,335	168,279	103,376	630,677	228.9%	557,091	1,085	49989.0%	50047.6%	(73.5%)	50198.2%	78580.2%
09JUL	95,709	699,358	54.2%	868,021	169,895	87,248	649,261	261.0%	557,178	1,156	50009.8%	50033.6%	(79.2%)	50208.3%	75722.4%
09AUG	95,063	692,639	50.7%	860,257	164,960	89,796	625,758	210.9%	559,615	1,033	50005.9%	50045.2%	(76.0%)	50238.1%	80755.5%
09SEP	95,386	723,565	45.3%	993,835	173,731	86,235	659,322	271.4%	558,695	1,209	50010.7%	50040.3%	(83.3%)	50273.1%	74333.1%
09OCT	110,516	751,360	40.5%	1,120,165	170,894	89,790	677,675	343.8%	551,679	1,115	50023.2%	50041.5%	(88.2%)	50320.0%	77606.3%
09NOV	101,695	790,149	84.7%	842,532	156,284	88,420	650,495	315.2%	547,744	1,033	50015.1%	50092.8%	(73.1%)	50195.9%	79127.8%
09DEC	93,853	708,598	55.3%	877,259	195,606	91,341	612,324	232.8%	548,255	1,182	50002.8%	50085.7%	(76.2%)	50292.9%	78507.9%
10JAN	92,361	693,863	65.2%	797,518	143,821	96,678	652,702	48.4%	815,723	132,727	49995.5%	50027.0%	34.7%	49988.1%	50008.2%
10FEB	99,433	700,474	78.4%	755,709	134,556	88,638	665,805	52.9%	813,433	118,320	50012.2%	50020.9%	48.2%	49965.2%	50013.8%

Figura 4.4. Relatório 3 ROLAP no MicroStrategy.

4.5.1.4. Relatório 4.

Este relatório apresenta 4 atributos e 3 métricas.

Os atributos são o Year e o atributo Month, que pertencem à mesma dimensão (dimensão Date) e tendo o atributo Year apenas cinco elementos (2007, 2008, 2009, 2010 e 2011). Os outros dois atributos são um a que se chamará de ST e que tem 9 elementos e o outro será o CL que tem também ele 9 elementos.

As métricas serão a métrica A, uma métrica F e uma G. A métrica A já foi descrita anteriormente, a métrica F é uma métrica condicional e é também uma das métricas que fazem parte da métrica composta A. A métrica G é também uma métrica condicional cujo facto utiliza a função SUM sobre uma coluna da tabela de factos. A todas estas métricas é aplicado um filtro para excluir um elemento de um atributo que designaremos por TS e outro elemento do atributo IL.

Existem valores destas métricas para a maior parte das combinações destes 4 atributos.

4.5.1.5. Relatório 5.

Este relatório tem 5 atributos e 6 métricas. Os atributos são o Year e Month, que pertencem à mesma dimensão (dimensão Date), o atributo RL que tem 50 elementos, o atributo BB com 6 elementos e, por fim, o atributo CB com 4 elementos.

As métricas são a métrica H que é uma métrica condicional à qual é aplicado um filtro para excluir um elemento de um atributo que designaremos por TS e outro elemento do atributo IL. Esta métrica utiliza um facto com a função SUM sobre uma das colunas da tabela de factos.

A métrica I e a métrica J são as duas métricas que fazem parte da métrica composta E. Estas duas métricas utilizam um facto com uma função COUNT que se aplica a duas colunas de uma das tabelas de factos. A estas duas métricas é aplicado um filtro para excluir um elemento de um atributo que designaremos por TS e outro elemento do atributo IL.

As métricas que se seguem são as métricas E, o Homólogo da métrica E e a Variação do valor actual em relação ao Homólogo da métrica E.

A maior parte destas combinações entre os elementos dos atributos tem valores para todas as métricas.

4.5.2. Relatórios 6 e 7.

Os dois relatórios que se seguem diferem dos anteriores por terem filtros. Sabendo-se que nos relatórios podem ser aplicados *report filters* e *view filters*, então iremos ter um relatório para estas duas categorias de filtros. Um *report filter* é um filtro que “especifica as condições que os dados devem ter para serem incluídos nos resultados de um relatório” (MicroStrategy, 2009b), estes filtros aparecem no SQL Statement de um relatório. Um *view filter* é um filtro “em tempo real, com base apenas nos objectos que existem no relatório, sendo aplicado apenas ao conjunto de dados devolvidos do *data warehouse*” (MicroStrategy, 2009b). Não aparece no SQL Statement de um relatório. Com estes relatórios será interessante analisar se quando aplicado um filtro existem diferenças entre o ROLAP e MOLAP e, se se verificarem, analisar se acontecem com ambos os filtros.

4.5.2.1. Relatório 6.

Tem apenas um atributo, o Month, e uma métrica, a métrica B. Este relatório tem ainda um *report filter*.

4.5.2.2. Relatório 7.

Este relatório é praticamente igual ao anterior, com apenas um atributo, o Month, e uma métrica, a métrica B. No entanto, difere do relatório anterior no filtro, pois não tem um *report filter*, mas sim um *view filter*.

Month	Metric B
10JAN	693.863
10FEB	700.474
10MAR	681.485
10APR	661.770
10MAY	704.665
10JUN	731.790
10JUL	711.622
10AUG	685.622
10SEP	719.049
10OCT	758.518
10NOV	725.170
10DEC	708.215

Figura 4.5. Relatório 7 MOLAP no MicroStrategy.

4.5.3. Relatórios 8 a 10.

Os próximos três relatórios tiveram como objectivo cada um deles reunir um dos seguintes tipos de métricas: compostas, condicionais e de transformação. À semelhança dos relatórios anteriores, também com estes pretendemos comparar se existem diferenças entre os relatórios das tabelas e do cubo.

4.5.3.1. Relatório 8.

Este relatório é composto pelo atributo Month e por 5 métricas: a métrica E, A, C, D e L, sendo esta última uma métrica composta por duas métricas condicionais que, por sua vez, utilizam um facto com a função SUM sobre uma coluna da tabela de factos. A estas métricas condicionais é aplicado um filtro para excluir um elemento de um atributo que designaremos por TS e outro elemento do atributo IL.

Estas 5 métricas têm em comum o facto de serem métricas compostas.

4.5.3.2. Relatório 9.

Este relatório é composto por um atributo, o atributo Month, e 4 métricas: a métrica F, G, B e uma nova métrica, que designaremos por métrica M. Esta métrica M utiliza um facto com a função COUNT sobre uma coluna de uma tabela de factos. A esta métrica M é aplicado um filtro para excluir um elemento de um atributo que designaremos por TS e outro elemento do atributo IL.

Estas quatro métricas têm em comum o facto de serem métricas condicionais.

Month	Metric B	Metric F	Metric G	Metric M
08JAN	660.022	146,849,015	91,673	
08FEB	648.211	219,368,366	78,601	
08MAR	656.676	47,010,028	85,740	
08APR	657.231	74,183,526	99,896	
08MAY	659.114	176,419,751	97,874	
08JUN	630.677	56,234,257	104,876	
08JUL	649.261	56,003,103	88,748	
08AUG	625.758	58,515,456	91,296	
08SEP	659.322	57,605,054	87,735	
08OCT	677.675	50,690,675	91,290	
08NOV	650.495	46,736,278	89,911	2,009
08DEC	612.324	46,929,384	92,545	2,296
09JAN	652.702	64,741,699	97,048	3,130
09FEB	665.805	128,776,472	89,641	2,497
09MAR	777.714	671,452,352	103,409	2,090
09APR	777.155	814,353,535	110,200	
09MAY	719.573	72,264,300	91,979	2,001
09JUN	692.848	52,729,690	93,515	
09JUL	699.358	66,685,895	97,209	
09AUG	682.639	62,281,633	96,563	
09SEP	723.565	214,992,501	96,886	
09OCT	751.360	356,298,104	112,016	
09NOV	790.149	111,458,561	102,857	2,328
09DEC	708.589	65,145,040	95,092	2,261
10JAN	693.863	80,107,598	93,861	
10FEB	700.474	63,050,274	100,933	

Figura 4.6. Relatório 9 ROLAP no MicroStrategy.

4.5.3.3. Relatório 10.

Este relatório é composto por um atributo, o atributo Month, e 5 métricas. Estas métricas são os Homólogos, i.e., as métricas que retornam os valores do mês homólogo das métricas E, A, C, D e B. Estas métricas têm em comum serem métricas de transformação.

4.5.4. Relatórios 11 e 12.

Os dois relatórios que se seguem diferem dos anteriores por terem um *prompt*, i.e., “um objecto utilizado nos relatórios que pergunta ao utilizador, durante a fase de execução do relatório”, qual a informação que pretende que lhe seja retornada de entre um conjunto pré-definido. (MicroStrategy, 2009b) Existem quatro tipos de *prompts*: *filter definition prompt*, *object prompt*, *value prompt* e *level prompt*.

O *filter definition prompt* permite “seleccionar critérios de filtragem das hierarquias, atributos, listas de elementos de atributo e métricas” (MicroStrategy, 2009b). Já o *object prompt* permite “seleccionar os objectos a incluir num relatório”, enquanto o *value prompt* permite “seleccionar um único valor sobre o qual se qualifica” e o *level prompt* permite “especificar o nível para uma métrica” (MicroStrategy, 2009b). Destes filtros só se irá utilizar o *filter definition prompt*. Isto porque dentro dos quatro tipos de *prompt* existem ainda subtipos, o que faz com que exista uma grande variedade de *prompts*. Juntamente com a grande variedade de *prompts* existentes, o facto de os *prompts* serem filtros dinâmicos e de já existirem relatórios com filtros que se irão comparar, levou a que se escolhesse comparar relatórios ROLAP com relatórios MOLAP para apenas dois tipos de *prompts*. Neste caso, a escolha recaiu sobre *filter definition prompts*: um será um *attribute prompt*, i.e., uma lista de todos os elementos do atributo escolhido, e o outro será um *attribute element prompt*, em que se selecciona uma lista pré-definida dos elementos. A escolha sobre estes dois tipos de *filter definition prompt* deveu-se ao facto de normalmente serem muito utilizados.

4.5.4.1. Relatório 11.

Este relatório tem 3 atributos e 2 métricas. Os atributos são o RL e o RC, que pertencem à mesma dimensão, e o BB. As métricas são a E e a H. Este relatório tem também um *prompt*: um *filter definition prompt* sobre o atributo RL. No caso deste relatório o filtro será uma lista de todos os elementos do atributo escolhido, neste caso o RL.

4.5.4.2. Relatório 12.

Este relatório é igual ao anterior, tendo os mesmos atributos e métricas, diferindo apenas no *prompt* utilizado. É também um *filter definition prompt*, mas neste caso em vez de uma lista de todos os elementos do atributo ou de se escolher utilizar um *prompt* criado à parte, irá seleccionar-se uma lista pré-definida dos elementos.

4.5.5. Relatório 13.

O relatório que se segue difere dos anteriores por ter um filtro que retorna só os 9 elementos de um atributo que têm o maior valor para uma métrica. Também com este relatório, tal como aconteceu com os anteriores, se pretende ver se existem diferenças entre o relatório com dados das tabelas e o relatório com dados do cubo.

4.5.5.1. Relatório 13.

Este relatório tem apenas o atributo AC, que possui cerca de 230 elementos, e a métrica N que é uma métrica condicional cujo facto é um Count de uma das colunas da tabela de factos e à qual se aplica um filtro para excluir um elemento do atributo IL. O que torna este relatório diferente dos anteriores é que tem um filtro (um filtro sobre uma função, neste caso, a função Rank) criado para retornar só os 9 elementos do atributo AC com base nos valores da métrica N (i.e., os 9 elementos que têm valores mais elevados para a métrica N).

Relatório 13_MOLAP

File Edit View Insert Format Data Grid Move Window Help

Save and Close Corporate Select... Values

Report objects

Name	Type
AC	Attribute
Metric N	Metric

Report details

Report Filter:
{Top 9 Countries} = 1

Report Cache Used: No

View filter

	Metrics	Metric N
AC		
SWITZERLAND		12,949
U.S. OF AMERICA		13,567
UNITED KINGDOM		18,285
PORTUGAL		242,979
SPAIN		15,445
NETHERLANDS		8,919
ITALY		16,058
BRAZIL		95,126
FRANCE		19,505

Figura 4.7. Relatório 13 MOLAP no MicroStrategy.

4.5.6. Relatório 14.

Este relatório distingue-se dos já apresentados por possuir um maior número de atributos, sendo por isso um relatório maior do que os anteriores.

4.5.6.1. Relatório 14.

Neste relatório encontram-se 4 métricas e 6 atributos. As métricas já foram todas apresentadas anteriormente, sendo elas as métricas E, I, J e N. Os atributos são todos de hierarquias diferentes: Month, IL, AC, RL, CL e BB. Alguns destes atributos têm um

grande número de elementos, como é o caso do AC e do RL, o que tornará este relatório maior do que os anteriores.

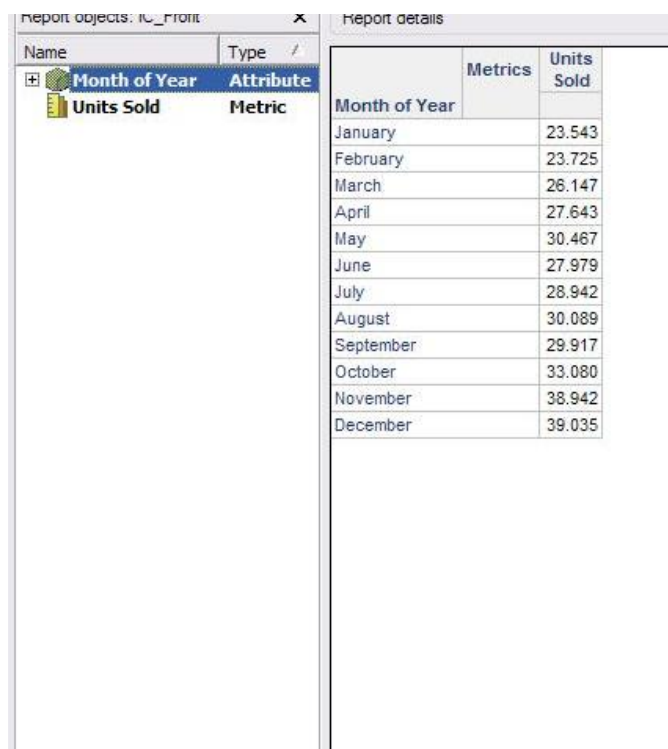
4.5.7. Relatórios 15 a 18.

Os próximos quatro relatórios estão relacionados. Um deles não é um relatório, mas sim um *Intelligent Cube*. Um *Intelligent Cube* é um “cubo multidimensional que pode ser compartilhado como uma única cópia na memória, entre muitos relatórios diferentes criados por vários utilizadores”. (MicroStrategy, 2009a) Sobre este *Intelligent Cube* são criados os próximos três relatórios.

O facto de estes relatórios serem construídos sobre um *Intelligent Cube* faz com que sejam diferentes dos apresentados anteriormente, pois os *Intelligent Cube* são uma combinação do relacional e do multidimensional, i. e., “combinam a rapidez e interactividade das análises MOLAP com o poder analítico e a profundidade do ROLAP” (MicroStrategy, 2011). Estes relatórios em vez de retornarem os dados do *data warehouse*, retornam os dados que se encontram no *Intelligent Cube* e que estão guardados na memória do Intelligent Server (MicroStrategy, 2009a). Por isso, uma vez publicado o *Intelligent Cube*, os relatórios criados sobre ele retornam os dados instantaneamente, sem correr uma consulta no *data warehouse*, daí não fazer sentido analisar o tempo que estes demoram a correr, mas apenas o espaço que ocupam. Os *Intelligent Cubes* podem ser criados sobre duas fontes de dados: tabelas e cubos. A decisão de comparar relatórios criados sobre um *Intelligent Cube* (que é uma combinação de ROLAP e MOLAP) deveu-se apenas a pretender verificar as diferenças existentes (só se poderá verificar ao nível do tamanho da cache) entre estes relatórios criados sobre um *Intelligent Cube* e os restantes relatórios ROLAP e MOLAP.

Nestes três relatórios criados sobre o *Intelligent Cube* irá utilizar-se um *derived element*, que é um “agrupamento dos elementos de um atributo num relatório cujos dados retornam de um *Intelligent Cube* e, por isso o relatório não precisa de gerar ou re-

executar o SQL” (MicroStrategy, 2009a). Este agrupamento pode ser útil tanto para questões de análise como de formatação, uma vez que proporciona uma nova visão dos dados presentes no relatório (MicroStrategy, 2009a). Para se perceber melhor o que é um *derived element*, apresenta-se de seguida um exemplo que irá ajudar a explicá-lo.

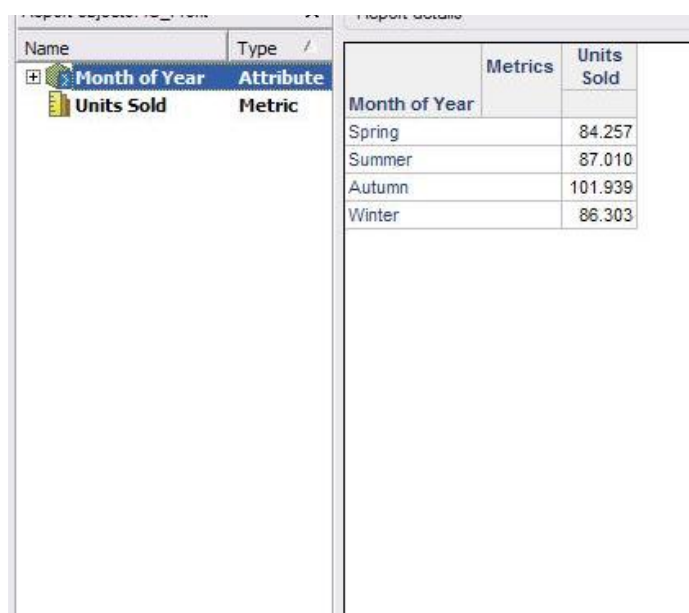


Month of Year	Units Sold
January	23.543
February	23.725
March	26.147
April	27.643
May	30.467
June	27.979
July	28.942
August	30.089
September	29.917
October	33.080
November	38.942
December	39.035

Figura 4.8. Exemplo de um relatório simples no MicroStrategy.

Na Figura 4.8. encontra-se um relatório com os dados de uma base de dados disponibilizada com o software MicroStrategy. Neste relatório, criado também ele sobre um *Intelligent Cube*, o atributo Month apresenta os valores da métrica “Units Solds” para cada mês.

Na Figura 4.9, num relatório igual ao visualizado anteriormente é utilizado um *derived element* sobre o atributo Month. Este *derived element* é utilizado para agrupar os meses por estações do ano (*Spring* agrupa os meses de *March*, *April* e *May*; *Summer* agrupa os meses de *June*, *July* e *August*; *Autumn* os meses de *September*, *October* e *November* e, por fim, *Winter* que agrupa os meses de *December*, *January* e *February*).



The screenshot shows the MicroStrategy report designer interface. On the left, a 'Report Explorer' pane displays a hierarchy with 'Month of Year' as an 'Attribute' and 'Units Sold' as a 'Metric'. On the right, a 'Report Outline' pane shows a table with the following data:

Month of Year	Units Sold
Spring	84.257
Summer	87.010
Autumn	101.939
Winter	86.303

Figura 4.9. Exemplo de um relatório com um *derived element* no MicroStrategy.

Terminado este exemplo de um *derived element*, importa ainda referir que estes podem ser definidos de três formas: lista, filtro e cálculo. Cada um dos três relatórios apresentados corresponde a cada uma das diferentes formas dos *derived element*.

4.5.7.1. Relatório 15.

Este *Intelligent Cube* tem 2 atributos - o RL e o ILD, que é um atributo com mais detalhe que o IL apresentado já em anteriores relatórios. Tem também 2 métricas: a métrica H e a métrica I.

4.5.7.2. Relatório 16.

Este relatório tem dois atributos (o atributo RL e o ILD) e duas métricas (métricas H e I), estando-se a utilizar o atributo ILD para criar um *derived element*. Neste relatório estamos a utilizar o *derived element* definido por um cálculo.

4.5.7.3. Relatório 17.

Tal como o relatório anterior, este também é construído sobre um *Intelligent Cube*. Tem também dois atributos (o atributo RL e o ILD) e duas métricas (métricas H e I). À semelhança do relatório anterior, também o atributo ILD presente neste relatório utiliza um *derived element*, mas desta vez definido por um filtro.

4.5.7.4. Relatório 18.

Neste relatório, construído sobre um *Intelligent Cube*, também ele composto por dois atributos (atributos RL e ILD) e duas métricas (métricas H e I), utiliza-se um *derived element* definido por uma lista.

Estes 18 relatórios que se irão comparar foram escolhidos por terem objectos e características que eram possíveis serem realizados tanto em relatórios cuja fonte de dados é as tabelas como sobre relatórios em que a fonte de dados é o cubo.

De seguida iremos apresentar uma tabela com a informação resumo de cada relatório:

Quadro 4.1. *Apresentação dos Relatórios*

Relatório	Métricas	Atributos	Características/Especificidades
Relatório 1	1 métrica (métrica A)	1 atributo (GP)	
Relatório 2	1 métrica (métrica A)	2 atributos (GP e Month)	
Relatório 3	15 métricas (métricas para o actual, homólogo e variação das métricas A, B, C, D, E)	1 atributo (Month)	
Relatório 4	3 métricas (métricas A, F, G)	4 atributos (Year, Month, ST, CL)	
Relatório 5	6 métricas (métricas H, I, J, E Actual, E Homólogo e E Variação)	5 atributos (Year, Month, RL, BB, CB)	
Relatório 6	1 métrica (métrica B)	1 atributo (Month)	Tem um <i>report filter</i> .
Relatório 7	1 métrica (métrica B)	1 atributo (Month)	Tem um <i>view filter</i> .
Relatório 8	5 métricas (métricas E, A, C, D e L)	1 atributo (Month)	Só tem métricas compostas.
Relatório 9	4 métricas (métricas F, G, B e M)	1 atributo (Month)	Só tem métricas condicionais.
Relatório 10	5 métricas (Homólogo das métricas E, A, C, D e B)	1 atributo (Month)	Só tem métricas de transformação.
Relatório 11	2 métricas (métricas E e H)	3 atributos (RL, RC, BB)	Tem um <i>attribute prompt</i> (lista de todos os elementos do atributo).
Relatório 12	2 métricas (métricas E e H)	3 atributos (RL, RC, BB)	Tem um <i>attribute element prompt</i> (lista pré-definida dos elementos).

Relatório 13	1 métrica (métrica N)	1 atributo (AC)	Tem um filtro dos 9 com maior valor.
Relatório 14	4 métricas (métricas E, I, J e N)	6 atributos (Month, IL, AC, RL, CL, BB)	
Relatório 15	2 métricas (métricas H e I)	2 atributos (RL, ILD)	<i>Intelligent Cube.</i>
Relatório 16	2 métricas (métricas H e I)	2 atributos (RL, ILD)	Utiliza-se um <i>derived element</i> definido por cálculo.
Relatório 17	2 métricas (métricas H e I)	2 atributos (RL, ILD)	Utiliza-se um <i>derived element</i> definido por um filtro.
Relatório 18	2 métricas (métricas H e I)	2 atributos (RL, ILD)	Utiliza-se um <i>derived element</i> definido por uma lista.

4.6. Limitações ROLAP e MOLAP

Após a apresentação dos relatórios a utilizar para comparar o ROLAP com o MOLAP, importa também abordar as limitações de cada um destes tipos de fontes de dados na criação dos relatórios.

Ambos os tipos de fontes de dados apresentam como limitação não conseguirem criar *derived elements* em relatórios. Estes só podem ser criados sobre *Intelligent Cubes*, i. e., fontes de dados que combinam o ROLAP com o MOLAP. Esta situação veio limitar a “criação de grupos na altura (*on-the-fly*) da visualização de um relatório” (MicroStrategy, 2009a) para propósitos de análise e formatação.

Para suprimir esta limitação, no ROLAP pode-se utilizar *custom groups* ou *consolidations*. Contudo, estas suas soluções não se encontram disponíveis *on-the-fly* tal como acontecia com os *derived elements*. Os *custom groups* são “objectos compostos por uma colecção ordenada de elementos em que cada elemento contém o seu próprio conjunto de qualificações de filtragem” (MicroStrategy, 2010), enquanto as *consolidations* são também “objectos compostos por uma colecção ordenada de

elementos”, mas neste caso “cada elemento é um agrupamento de elementos de um atributo que fornece operações aritméticas entre linhas” (MicroStrategy, 2010b). No entanto, estes dois objectos não se encontram disponíveis no MOLAP (MicroStrategy, 2010b).

Outra limitação encontrada nos relatórios MOLAP é o facto de não ser possível criar *stand-alone filters*. *Stand-alone filters* são “filtros criados como objectos MicroStrategy independentes” (MicroStrategy, 2010a) e têm como vantagem que o mesmo filtro pode ser aplicado a vários relatórios, métricas e outros objectos. No ROLAP esta limitação não se verifica.

Apesar das métricas criadas no MOLAP (métricas com MDX) combinarem um conjunto de expressões e funções MDX com ferramentas analíticas MicroStrategy, neste tipo de fonte de dados a criação de métricas com definições complexas é mais difícil (MicroStrategy, 2010b).

5. APRESENTAÇÃO DOS RESULTADOS

Após terem-se executado 5 vezes cada um dos relatórios explicados anteriormente, apresentam-se agora, no Quadro 5.1., os resultados da média aritmética do tempo de execução e da cache de cada um deles, de forma a facilitar a comparação dos relatórios ROLAP com os MOLAP.

Quadro 5.1. *Comparação dos Resultados dos Relatórios*

		Média	
		ROLAP	MOLAP
Relatório 1	Total Duration in Query Engine	0:00:03.57	0:00:07.13
	Cache (KB)	65,6	38,4
Relatório 2	Total Duration in Query Engine	0:00:03.13	0:00:05.33
	Cache (KB)	103	77
Relatório 3	Total Duration in Query Engine	0:00:14.84	0:00:05.31
	Cache (KB)	259,4	141
Relatório 4	Total Duration in Query Engine	0:00:03.44	0:00:05.14
	Cache (KB)	142,2	83
Relatório 5	Total Duration in Query Engine	0:15:02.14	0:00:14.70
	Cache (KB)	798	672,2
Relatório 6	Total Duration in Query Engine	0:00:01.21	0:00:04.98
	Cache (KB)	50	39
Relatório 7	Total Duration in Query Engine	0:00:01.06	0:00:05.62
	Cache (KB)	60	51
Relatório 8	Total Duration in Query Engine	0:00:06.60	0:00:04.66
	Cache (KB)	175,4	66
Relatório 9	Total Duration in Query Engine	0:00:03.65	0:00:05.28
	Cache (KB)	87	50,4
Relatório 10	Total Duration in Query Engine	0:00:08.63	0:00:04.83
	Cache (KB)	209,4	60
Relatório 11	Total Duration in Query Engine	0:00:01.74	0:00:04.67
	Cache (KB)	81,6	46,4
Relatório 12	Total Duration in Query Engine	0:00:01.23	0:00:04.83
	Cache (KB)	127	96
Relatório 13	Total Duration in Query Engine	0:01:12.47	0:00:04.93
	Cache (KB)	53,2	36

Relatório 14	Total Duration in Query Engine	0:22:38.88	-
	Cache (KB)	13862	-
Relatório 15	Total Duration in Query Engine	0:00:00.07	0:00:00.05
	Cache (KB)	253,6	14772,2
Relatório 16	Total Duration in Query Engine	-	-
	Cache (KB)	73,2	70,2
Relatório 17	Total Duration in Query Engine	-	-
	Cache (KB)	72,4	71,2
Relatório 18	Total Duration in Query Engine	-	-
	Cache (KB)	71,6	49,6

De seguida apresentam-se os resultados do Quadro 5.1. graficamente: na Figura 5.1.encontra-se a média aritmética do tempo que os relatórios demoraram a executar e na Figura 5.2. a cache de cada um deles.

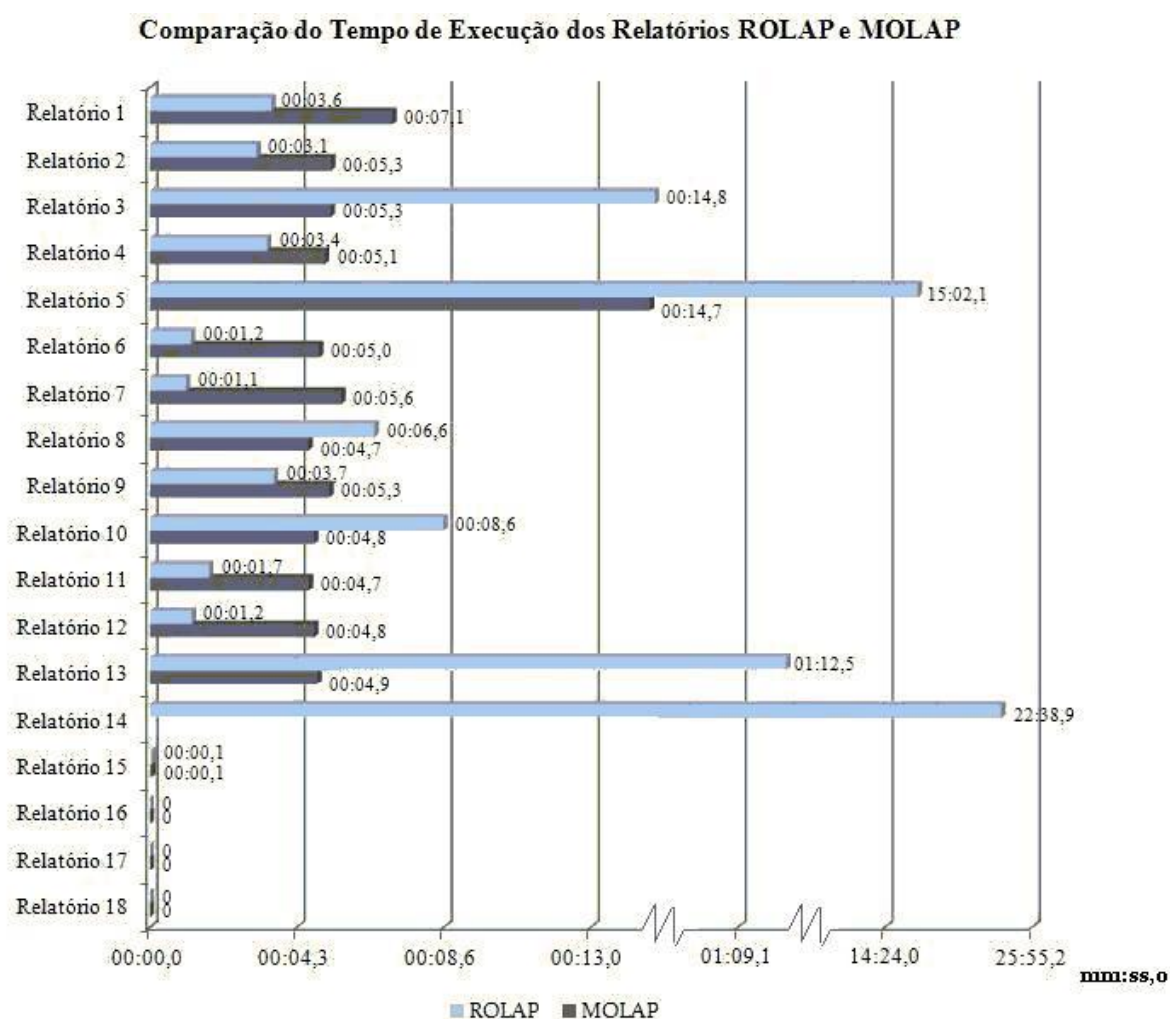


Figura 5.1. Comparação do tempo de execução entre relatórios ROLAP e MOLAP.

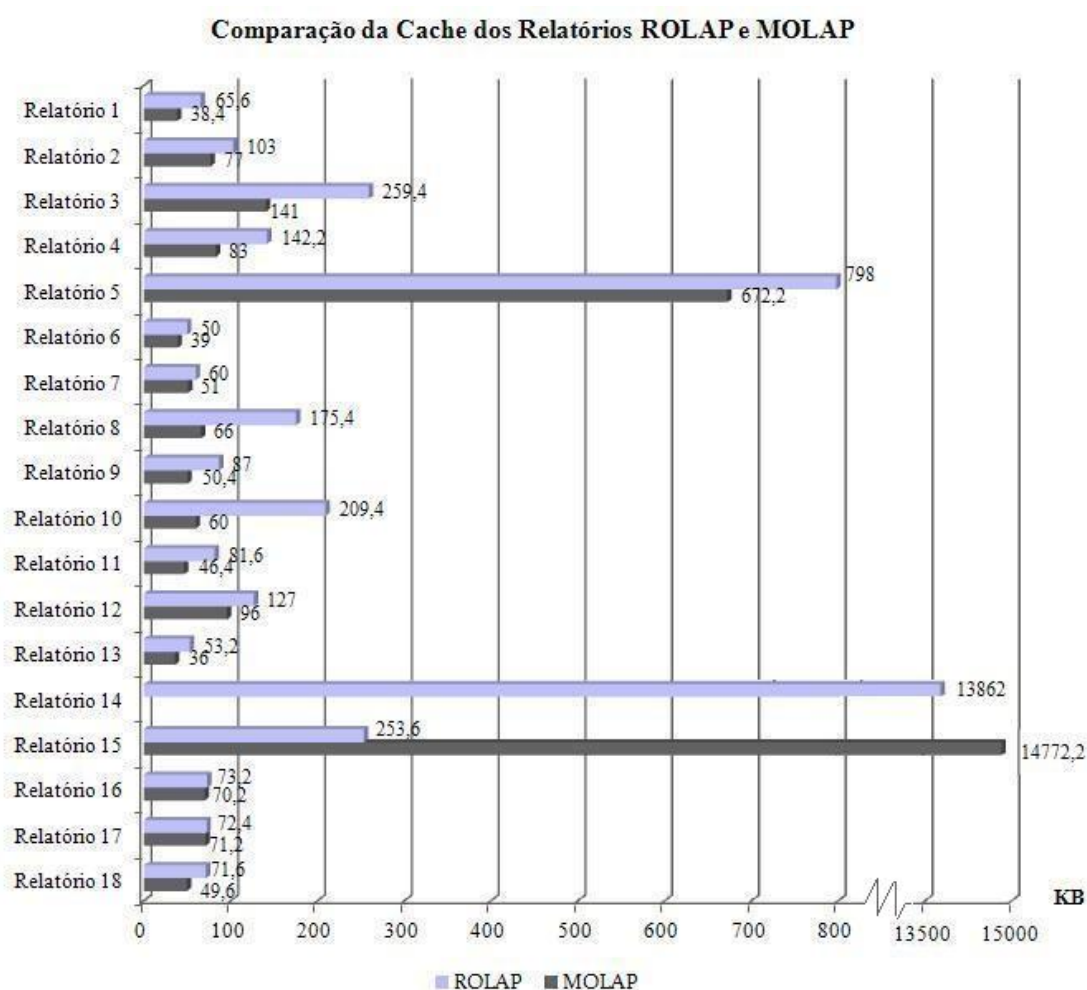


Figura 5.2. Comparação da cache entre relatórios ROLAP e MOLAP.

Para uma análise mais pormenorizada, no apêndice A encontram-se duas tabelas (uma para o ROLAP e outra para o MOLAP) com os resultados do tempo de execução e da *cache* de cada um dos relatórios nas 5 vezes em que foram executados. De seguida, irá fazer-se uma breve descrição dos resultados.

Relativamente ao primeiro relatório, que é composto só por um atributo e uma métrica, o tempo de execução para o ROLAP foi aproximado nas cinco vezes, enquanto no caso do MOLAP registou-se na primeira vez um tempo de execução demasiado elevado comparativamente ao das seguintes vezes.

As *caches* deste relatório registam o mesmo tamanho ao longo das três primeiras execuções, aumentando nas duas últimas execuções do relatório MOLAP e diminuindo

no relatório ROLAP. No entanto, verifica-se que o relatório MOLAP ocupa menos espaço que o ROLAP.

Analisando a médias das cinco execuções conclui-se que este relatório MOLAP demora mais tempo a ser executado do que o ROLAP, mas por sua vez o MOLAP tem quase metade do tamanho da *cache* do ROLAP.

O segundo relatório tem apenas mais um atributo que o anterior. O tempo de execução dos relatórios nas cinco vezes é semelhante. À semelhança do que aconteceu com o relatório anterior, a *cache* do relatório MOLAP é inferior à do ROLAP.

Relativamente à média deste relatório verifica-se que o MOLAP apresenta quase o dobro do tempo de execução, mas por sua vez o ROLAP também apresenta um valor bastante superior de *cache*.

Ao contrário do que seria de esperar, uma vez que este relatório tem apenas mais um atributo em relação ao Relatório 1, tanto o ROLAP como o MOLAP demoram menos tempo a serem executados do que o Relatório 1.

O Relatório 3 tem 15 métricas, calculando 5 delas o valor Homólogo e outras 5 a Variação do valor actual em relação ao Homólogo. Neste caso o tempo de execução registado é semelhante entre os relatórios ROLAP e também entre os relatórios MOLAP. O relatório ROLAP demora, em média, cerca de três vezes mais tempo que o relatório MOLAP. As *caches* do relatório ROLAP continuam a ser maiores do que a do MOLAP.

O quarto relatório tem quatro atributos e três métricas e os valores registados em todas as execuções são muito próximos. O relatório MOLAP continua, em média, a demorar mais tempo a ser executado, já a *cache* continua a ser maior no caso do relatório ROLAP.

O Relatório 5 é composto por cinco atributos e seis métricas. Os valores registados durante as cinco execuções não são muito diferentes entre eles.

Tal como tinha acontecido com o Relatório 3, também neste caso o relatório ROLAP demorou muito mais tempo a ser executado do que o relatório MOLAP. A diferença do tempo demorado entre o ROLAP e o MOLAP é muito grande, tendo o ROLAP demorado, em média, mais de 14 minutos a ser executado, enquanto o MOLAP, em média, demorou menos de 1 minuto.

Neste caso, a *cache* dos relatórios tem um comportamento semelhante à dos anteriores, verificando-se que a *cache* do ROLAP é superior à do MOLAP.

O Relatório 6 é composto por um atributo, uma métrica e um *report filter*. Das vezes que se executaram os relatórios, estes apresentaram valores semelhantes.

O relatório MOLAP continua a ter um maior tempo de execução do que o relatório ROLAP e continua também a ter uma *cache* de menor tamanho.

Passando ao Relatório 7, este é igual ao anterior com a diferença de que não tem um *reprot filter*, mas sim um *view filter*. Neste relatório o tempo de execução do relatório MOLAP continua a ser superior ao do ROLAP e a *cache* deste continua também a ser mais pequena do que a do ROLAP.

Comparando estes valores com os do Relatório 6 conclui-se que os tempos de execução são próximos entre um relatório com o *report filter* e outro com o *view filter*, mas tal como era de esperar a *cache* dos relatórios com o *view filter* é maior do que a dos que tem o *report filter*.

No Relatório 8, que é composto por um atributo e cinco métricas compostas, verifica-se uma situação curiosa. Isto porque das duas primeiras vezes que se correu o relatório ROLAP obteve-se exactamente o mesmo tempo de execução e o mesmo aconteceu com o relatório MOLAP, mas neste caso quando se correu pela primeira e pela última vez.

Também neste relatório se verifica uma alteração do tamanho da *cache* nas cinco vezes que se correu.

Comparando o tempo de execução, em média, entre o relatório ROLAP e MOLAP verifica-se que desta vez o ROLAP demora mais tempo e tem também uma *cache* muito maior que a do MOLAP.

Com um atributo e quatro métricas condicionais, o Relatório 9 à semelhança do anterior também apresenta um tempo de execução igual entre as duas últimas vezes e depois entre as duas últimas em que se executou o relatório MOLAP.

Em média, o tempo de execução do relatório MOLAP é superior ao do ROLAP, tal como a *cache* do ROLAP é superior à do MOLAP.

No Relatório 10, que é composto por um atributo e cinco métricas que utilizam uma transformação, o tempo de execução é semelhante entre eles das vezes que foram executados.

O tempo de execução do relatório MOLAP é mais uma vez superior ao do relatório ROLAP, sendo também as *caches* do relatório ROLAP maiores do que as do MOLAP.

Em relação ao Relatório 11, este tem três atributos, duas métricas e um *attribute prompt* cujos elementos escolhidos foram sempre os mesmos para todas as execuções.

Analisando os tempos de execução deste relatório ROLAP repara-se que numa das vezes em que foi executado existe uma grande diferença em relação aos tempos verificados nas outras vezes.

Nesta caso a diferença entre o tempo de execução, em média, entre um relatório ROLAP e um MOLAP é mais do dobro (demorando o MOLAP mais tempo), enquanto o tamanho das *caches* do ROLAP também são quase o dobro do tamanho das *caches* do MOLAP.

O Relatório 12 é igual ao anterior, diferindo apenas no *prompt* utilizado, que neste caso é um *attribute element prompt* cujos elementos escolhidos foram sempre iguais para todas as execuções. Os tempos de execução são próximos entre as cinco vezes que foi executado. Também sem surpresa são o tempo de execução (relatório MOLAP demora mais que o ROLAP) e o tamanho da *cache* (relatório ROLAP continua a ter uma *cache* maior que o MOLAP).

Relativamente ao Relatório 13, que é composto apenas por um atributo, uma métrica e um filtro dos 9 elementos com maior valor, os resultados obtidos revelam que o ROLAP tem, em média, um tempo de execução superior ao MOLAP e que a sua *cache* também é maior que a do MOLAP.

O Relatório 14, que possui 4 métricas e 6 atributos, tendo os seus atributos um grande número de elementos comparativamente aos restantes atributos utilizados noutros relatórios, o que faz dele um relatório maior que os restantes. Se no caso do ROLAP o seu tempo de execução foi, em média, cerca de 22 minutos, no caso do MOLAP ao final de 1 hora ainda não tinha terminado, dando um erro de memória insuficiente. Por isso, não se pode comparar o tempo de execução nem a *cache* do ROLAP com o MOLAP para este relatório, uma vez que o MOLAP não chegou a terminar de correr em nenhuma das cinco vezes. Pode-se apenas concluir que o relatório MOLAP demora pelo menos mais do triplo do tempo que o ROLAP.

No caso do Relatório 15, neste caso não é um relatório, mas sim um *Intelligent Cube* com dois atributos e duas métricas. O tempo de execução é igual tanto no ROLAP como no MOLAP, no entanto encontra-se uma grande diferença entre a *cache* de um e a do outro. Ao contrário do que tem sido habitual nos relatórios anteriores, neste caso o *Intelligent Cube* MOLAP tem uma *cache* bastante superior à do ROLAP.

No Relatório 16 está-se perante um relatório com dois atributos, duas métricas e com um *derived element* definido por um cálculo. Como tal, não se pode analisar o tempo de execução. Mas observando-se o espaço da *cache* conclui-se que a *cache* do relatório ROLAP continua a ser maior, mesmo que apenas ligeiramente, que a do relatório MOLAP.

Quanto ao Relatório 17 e à semelhança do relatório anterior, este também tem dois atributos e duas métricas. Tem também um *derived element* definido por um filtro. Também neste caso, a *cache* do relatório ROLAP continua a ser maior do que a do MOLAP, mas comparando com a *cache* do Relatório 16, o relatório MOLAP aumentou a sua ligeiramente.

O último relatório (Relatório 18) é igual aos dois anteriores, diferenciando-se apenas no *derived element* que, neste caso, é definido por uma lista. Esta diferença tem impacto na *cache* dos mesmos, que reduz para 50 KB no caso do relatório MOLAP. A *cache* do relatório MOLAP continua a ser inferior à do ROLAP.

6. DISCUSSÃO DOS RESULTADOS

Os resultados obtidos permitiram chegar a algumas conclusões. Analisando-os conclui-se que, neste caso e utilizando o software MicroStrategy, os relatórios MOLAP normalmente demoram mais tempo a executar do que os ROLAP. No entanto, verificaram-se as seguintes exceções:

1. Sempre que se utilizam métricas de transformação sobre o relacional (Homólogo e Variação).
2. No relatório onde só se encontravam métricas compostas. Apesar de existirem relatórios onde se encontram métricas compostas juntamente com outro tipo de métricas, só neste caso, onde unicamente existiam métricas compostas, se registou esta diminuição de performance do ROLAP em relação ao MOLAP.
3. Quando se utilizou um filtro com uma função que retornaria os 9 elementos de um atributo com maiores valores para uma determinada métrica.
4. Num *Intelligent Cube*.

Algumas destas conclusões são esperadas de acordo com o que se apresentou na revisão da literatura (capítulo 2). Colliat (1996) apresentava como uma das razões para a utilização do MOLAP em detrimento do ROLAP a rapidez no cálculo de dados derivados, o que se verificou nos Relatórios 3, 5 e 10 (relatórios MOLAP mais rápidos que os ROLAP). Nestes relatórios ROLAP existem métricas de transformação para calcular o Homólogo e a Variação, enquanto que nos relatórios MOLAP criaram-se métricas no cubo. A mesma justificação pode ser aplicada ao relatório só com métricas compostas (Relatório 8) e ao relatório que retornaria os 9 elementos com maior valor para uma determinada métrica (Relatório 13).

Outra conclusão já esperada era que os sistemas ROLAP eram mais adequados para grandes quantidades de dados, enquanto os sistemas MOLAP para pequenos volumes de dados (Pirnau & Botezatu, 2010; Kazi et al., 2010). Verifica-se esta situação no Relatório 14, relatório esse que era o maior dos 18 existentes, sendo o que iria retornar uma maior quantidade de dados. No ROLAP, este relatório foi executado,

demorando, em média, cerca de 22 minutos, já no MOLAP ao final de 1 hora a executar ainda não tinha terminado.

Com os resultados obtidos também se pode concluir que os relatórios MOLAP têm uma *cache* inferior à dos relatórios ROLAP. Em alguns casos esta diferença chega a ser quase o dobro. Mas tal como acontecia com o tempo de execução, também com o tamanho das *caches* existe uma excepção. Esta verifica-se no *Intelligent Cube*, em que o tamanho do *Intelligent Cube* criado sobre o cubo é muito superior ao do criado sobre o relacional.

No entanto, nos relatórios criados sobre o *Intelligent Cube* a diferença do tamanho da *cache* é quase nula, tendo sempre os relatórios do *Intelligent Cube* sobre o relacional uma *cache* superior.

Uma vez que o *Intelligent Cube* sobre o relacional tem um tempo de execução semelhante ao do criado sobre o cubo e a *cache* dos relatórios criados sobre os dois *Intelligent Cubes* também são semelhantes pode-se concluir que, neste caso, é preferível criar um *Intelligent Cube* sobre as tabelas. Além desta diferença na *cache* do *Intelligent Cube* cuja fonte de dados é o cubo e tendo em conta que os relatórios criados sobre estes retornam instantaneamente os dados, não existe mais nenhuma diferença em relação aos restantes relatórios ROLAP e MOLAP.

Esta conclusão de que o tamanho da *cache* dos relatórios sobre o MOLAP é inferior à dos relatórios sobre o ROLAP também já era esperada, uma vez que quando Colliat (1996) defendeu a utilização do MOLAP em detrimento do ROLAP, uma das razões que apresentou foi exactamente o facto de o MOLAP ocupar menos espaço. No entanto, Colliat (1996) apresentou ainda mais uma razão para defender a utilização do MOLAP: este era mais rápido que o ROLAP, o que não se verificou nestes relatórios. No entanto, não se pode esquecer que a ferramenta utilizada (MicroStrategy) está mais vocacionada para o ROLAP, o que deve ter influenciado os resultados.

7. CONCLUSÃO E RECOMENDAÇÕES

7.1. Conclusão

No início deste trabalho foram estabelecidas duas questões às quais se iria procurar responder. A primeira era:

1. Quais são as vantagens, desvantagens e limitações do ROLAP e do MOLAP?

Sobre esta questão há muito a referir. Até porque antes de se responder a esta, este trabalho permitiu conhecer dois termos: OLAP e OLTP. O OLTP suporta as operações diárias e utiliza principalmente sistemas de gestão de bases de dados relacionais. No entanto, o constante crescimento da quantidade de dados levou ao aparecimento do OLAP, de forma a colmatar as necessidades a que o OLTP não conseguia responder.

O OLAP suporta a análise de negócio, sendo um sistema de apoio à decisão. Com o aparecimento do OLAP surgiram vários tipos de OLAP, entre eles o ROLAP e o MOLAP.

O ROLAP armazena os dados numa base de dados relacional. No ROLAP existe uma boa standardização das tecnologias de base de dados. Podem manipular grandes quantidades de dados. Pode ter as funcionalidades de uma base de dados relacional e como acede directamente à base de dados relacional é mais fácil adicionar novas dimensões. Os dados estão sempre disponíveis. Permite pedidos *ad hoc* eficazmente.

A performance pode ser mais lenta e é limitada pelas funcionalidades SQL. Como acede directamente aos dados para realizar operações de *drill down* em tempo real é necessário muito poder de processamento e, consequentemente, hardware caro.

O MOLAP armazena os dados num cubo multidimensional. O MOLAP permite uma boa performance nas operações *slicing and dicing*, tendo um bom tempo de performance para uma consulta OLAP. Permite cálculos complexos.

Limitado na quantidade de dados a manipular. Não existe uma estandardização das tecnologias de base de dados. Realizar alterações como adicionar novas dimensões é complicado devido à falta de flexibilidade. Não permite pedidos *ad hoc* eficazmente.

As limitações existentes, neste caso aplicadas ao MicroStrategy, são várias, reflectindo-se principalmente sobre o MOLAP.

A segunda questão a que se procurou responder ao longo deste trabalho era:

2. Existem diferenças entre relatórios desenvolvidos no MicroStrategy utilizando fontes de dados ROLAP e MOLAP?

Os resultados obtidos revelam que ter uma fonte de dados ROLAP e outra MOLAP não é exactamente o mesmo. Existem diferenças entre elas.

Pode-se concluir que no MicroStrategy os relatórios sobre uma fonte de dados ROLAP demoram menos tempo a serem executados, mas também ocupam mais espaço que os relatórios sobre uma fonte de dados MOLAP. Posso também concluir que para grandes quantidades de dados deve-se utilizar o ROLAP e para o cálculo de dados derivados o MOLAP.

Conclui-se também que, no MicroStrategy, o ROLAP apresenta um melhor desempenho do que o MOLAP podendo-se confirmar o que se supunha: o MicroStrategy está mais vocacionado para o ROLAP.

7.1. Limitações

Uma limitação encontrada neste estudo é o facto de não existir a possibilidade de generalização destes resultados a outras ferramentas de *reporting*, uma vez que cada ferramenta é diferente e, no caso do MicroStrategy, está mais vocacionada para ROLAP do que MOLAP.

APÊNDICE A - Comparação dos Resultados dos Relatórios

Os relatórios ROLAP apresentaram os seguintes resultados:

Quadro 7.1. *Resultados dos Relatórios ROLAP*

		Execução 1	Execução 2	Execução 3	Execução 4	Execução 5
Relatório 1	Total Duration in Query Engine	0:00:03.27	0:00:04.66	0:00:03.43	0:00:02.93	0:00:03.58
	Cache (KB)	68	68	68	62	62
Relatório 2	Total Duration in Query Engine	0:00:03.10	0:00:02.90	0:00:02.91	0:00:03.51	0:00:03.23
	Cache (KB)	103	103	103	103	103
Relatório 3	Total Duration in Query Engine	0:00:16.34	0:00:17.32	0:00:14.00	0:00:13.16	0:00:13.40
	Cache (KB)	259	259	259	260	260
Relatório 4	Total Duration in Query Engine	0:00:03.13	0:00:03.21	0:00:03.18	0:00:04.10	0:00:03.57
	Cache (KB)	139	139	139	147	147
Relatório 5	Total Duration in Query Engine	0:17:06.86	0:17:03.93	0:16:37.50	0:11:29.02	0:11:33.40
	Cache (KB)	868	868	868	693	693
Relatório 6	Total Duration in Query Engine	0:00:01.13	0:00:01.42	0:00:01.13	0:00:01.03	0:00:01.34
	Cache (KB)	50	50	50	50	50
Relatório 7	Total Duration in Query Engine	0:00:01.23	0:00:01.09	0:00:01.13	0:00:01.38	0:00:01.06
	Cache (KB)	60	60	60	60	60
Relatório 8	Total Duration in Query Engine	0:00:06.73	0:00:06.73	0:00:06.08	0:00:07.28	0:00:06.20
	Cache (KB)	176	176	175	175	175

Relatório 9	Total Duration in Query Engine	0:00:03.54	0:00:03.21	0:00:04.44	0:00:03.49	0:00:03.58
	Cache (KB)	87	87	87	87	87
Relatório 10	Total Duration in Query Engine	0:00:09.09	0:00:08.33	0:00:08.48	0:00:08.93	0:00:08.31
	Cache (KB)	209	209	209	210	210
Relatório 11	Total Duration in Query Engine	0:00:01.37	0:00:01.13	0:00:04.72	0:00:00.71	0:00:00.78
	Cache (KB)	82	82	82	81	81
Relatório 12	Total Duration in Query Engine	0:00:01.01	0:00:01.76	0:00:00.88	0:00:01.18	0:00:01.31
	Cache (KB)	127	127	127	127	127
Relatório 13	Total Duration in Query Engine	0:01:18.57	0:01:04.63	0:00:59.92	0:00:39.45	0:00:39.78
	Cache (KB)	46	55	55	55	55
Relatório 14	Total Duration in Query Engine	0:22:59.29	0:20:17.27	0:25:21.12	0:22:50.96	0:24:35.76
	Cache	13862	13862	13862	13862	13862
Relatório 15	Total Duration in Query Engine	0:00:00.04	0:00:00.05	0:00:00.07	0:00:00.08	0:00:00.09
	Cache (KB)	254	254	254	253	253
Relatório 16	Total Duration in Query Engine					
	Cache (KB)	74	74	74	72	72
Relatório 17	Total Duration in Query Engine					
	Cache (KB)	72	72	72	73	73

Relatório 18	Total Duration in Query Engine					
	Cache (KB)	72	72	72	71	71

Já com os relatórios MOLAP obteve-se os seguintes resultados:

Quadro 7.2. *Resultados dos Relatórios MOLAP*

		Execução 1	Execução 2	Execução 3	Execução 4	Execução 5
Relatório 1	Total Duration in Query Engine	0:00:13.63	0:00:05.38	0:00:05.53	0:00:06.44	0:00:04.68
	Cache (KB)	38	38	38	39	39
Relatório 2	Total Duration in Query Engine	0:00:05.41	0:00:05.11	0:00:04.86	0:00:05.66	0:00:05.60
	Cache (KB)	77	77	77	77	77
Relatório 3	Total Duration in Query Engine	0:00:06.03	0:00:05.20	0:00:04.74	0:00:05.61	0:00:04.97
	Cache (KB)	141	141	141	141	141
Relatório 4	Total Duration in Query Engine	0:00:05.57	0:00:04.95	0:00:05.02	0:00:05.05	0:00:05.10
	Cache (KB)	93	93	93	68	68
Relatório 5	Total Duration in Query Engine	0:00:15.27	0:00:13.97	0:00:13.85	0:00:15.02	0:00:15.35
	Cache (KB)	653	653	653	701	701

Relatório 6	Total Duration in Query Engine	0:00:05.00	0:00:04.91	0:00:05.22	0:00:04.86	0:00:04.89
	Cache (KB)	39	39	39	39	39
Relatório 7	Total Duration in Query Engine	0:00:05.35	0:00:05.16	0:00:04.99	0:00:04.64	0:00:05.17
	Cache (KB)	51	51	51	51	51
Relatório 8	Total Duration in Query Engine	0:00:04.71	0:00:04.68	0:00:04.71	0:00:04.72	0:00:04.50
	Cache (KB)	74	64	64	64	64
Relatório 9	Total Duration in Query Engine	0:00:05.33	0:00:05.22	0:00:05.22	0:00:05.32	0:00:05.32
	Cache (KB)	50	50	50	51	51
Relatório 10	Total Duration in Query Engine	0:00:04.54	0:00:05.10	0:00:04.72	0:00:04.75	0:00:05.03
	Cache (KB)	60	60	60	60	60
Relatório 11	Total Duration in Query Engine	0:00:04.72	0:00:04.39	0:00:04.08	0:00:04.96	0:00:05.21
	Cache (KB)	46	46	46	47	47
Relatório 12	Total Duration in Query Engine	0:00:05.10	0:00:04.32	0:00:04.94	0:00:04.91	0:00:04.89
	Cache (KB)	96	96	96	96	96
Relatório 13	Total Duration in Query Engine	0:00:04.88	0:00:04.93	0:00:05.07	0:00:04.83	0:00:04.93
	Cache (KB)	36	36	36	36	36
Relatório 14	Total Duration in Query Engine					
	Cache (KB)					

Relatório 15	Total Duration in Query Engine	0:00:00.05	0:00:00.05	0:00:00.05	0:00:00.05	0:00:00.05
	Cache (KB)	14523	14523	14523	15146	15146
Relatório 16	Total Duration in Query Engine					
	Cache (KB)	71	71	71	69	69
Relatório 17	Total Duration in Query Engine					
	Cache (KB)	72	72	72	70	70
Relatório 18	Total Duration in Query Engine					
	Cache (KB)	50	50	50	49	49

REFERÊNCIAS

- Abelló, A., & Romero, O. (Eds.). (2009) *Encyclopedia of Database Systems*. Springer.
- Anzanello, C. A. (2002). *OLAP Conceitos e Utilização*.
- Asghar, S., Fong, S., & Rukhsana, A. (2009). *A Contemplation of Group Decision Support Systems*. Paper presented at the Proceedings of the 2009 Fourth International Conference on Computer Sciences and Convergence Information Technology.
- Asghar, S., & Osama, F. A. (2009). *A Literature Review on Efficient Online Analytical Processing System*. Kaunas: Kaunas Univ Technology Press.
- Baskerville, R. L. (1999). Investigating information systems with action research. *Commun. AIS*, 2.
- Bhan, M., Geetha, D. E., Kumar, T. V. S., Rajanikanth, K., & Society, I. C. (2009). *Modeling of Object Oriented OLAP*. New York: Ieee.
- Bispo, C. A. F. (1998). *Uma análise da nova geração de sistemas de apoio à decisão*. Escola de Engenharia de São Carlos da Universidade de São Paulo, São Carlos.
- Bog, A., Domschke, M., Mueller, J., & Zeier, A. (2009). A Framework for Simulating Combined OLTP and OLAP Workloads. In E. S. Qi, G. Cheng, J. A. Shen & R. L. Dou (Eds.), *2009 Ieee 16th International Conference on Industrial Engineering and Engineering Management, Vols 1 and 2, Proceedings* (pp. 1675-1678). New York: Ieee.
- Bog, A., Krueger, J., & Schaffner, J. (2008). *A Composite Benchmark for Online Transaction Processing and Operational Reporting*.
- Bog, A., Plattner, H., & Zeier, A. (2011). A mixed transaction processing and operational reporting benchmark. *Information Systems Frontiers*, 13(3), 321-335.

- Buzydlowski, J., Song, I., & Hassell, L. (1998). *A framework for object-oriented on-line analytic processing*. Paper presented at the Proceedings of the 1st ACM international workshop on Data warehousing and OLAP.
- Chaudhuri, S., & Dayal, U. (1997). An overview of data warehousing and OLAP technology. *SIGMOD Rec.*, 26(1), 65 - 74.
- Clay, C., Consulting, H. & Smith, B. (2009). *Microsoft SQL Server 2008 MDX Step by Step* Available from <http://www.wowebook.pro/book/microsoft-sql-server-2008-mdx-step-by-step/>
- Codd, F., Codd, S., & Salley, C. (1993). *Providing OLAP (On-line Analytical Processing) to User-Analysts: An IT Mandate*.
- Colliat, G. (1996). OLAP, Relational, and Multidimensional Database Systems. *SIGMOD Rec.*, 25(3), 64 - 69.
- Conn, S. (2005). *OLTP and OLAP data integration: A review of feasible implementation methods and architectures for real time data analysis*. New York: Ieee.
- Cramer, R. (2006). *Estudo Analítico de Ferramentas Open Source para Ambientes OLAP*. Universidade do Extremo Sul Catarinense - UNESC, Criciúma.
- Cuzzocrea, A. (2005). *Overcoming limitations of approximate query answering in OLAP*. Los Alamitos: Ieee Computer Soc.
- Datta, A., & Thomas, H. (1999). The cube data model: a conceptual model and algebra for on-line analytical processing in data warehouses. [doi: 10.1016/S0167-9236(99)00052-4]. *Decision Support Systems*, 27(3), 289-301.
- Datawarehouse4u.Info. (2008-2009). OLTP vs. OLAP. Retrieved 3rd November 2010: <http://datawarehouse4u.info/OLTP-vs-OLAP.html>
- DataWarehousingGuide.com. Data Warehousing Guide - Business Intelligence tutorials & tips. Retrieved 24th September 2011, from

<http://www.datawarehousingguide.com/microstrategy-interview-questions/11-basic-concepts-microstrategy-interview-question-and-answers-.html>

- Dehne, F., Eavis, T., & Rau-Chaplin, A. (2003). *Parallel multi-dimensional ROLAP indexing*. Los Alamitos: Ieee Computer Soc.
- Dehne, F., Eavis, T., & Rau-Chaplin, A. (2005). *Parallel querying of ROLAP cubes in the presence of hierarchies*. Paper presented at the Proceedings of the 8th ACM international workshop on Data warehousing and OLAP.
- Dehne, F., Eavis, T., & Rau-Chaplin, A. (2006). The cgmCUBE project: Optimizing parallel data cube generation for ROLAP. *Distributed and Parallel Databases*, 19(1), 29-62.
- Feiman, J., & MacDonald, N. (2011). *Magic Quadrant for Business Intelligence Platforms*: Gartner RAS Core Research.
- Goff, K. S., Langit, L., Malik, S., Mauri, D., Welch, J. (2009). *Smart Business Intelligence Solutions with Microsoft SQL Server 2008*: Microsoft Press.
- Gorla, N. (2003). Features to consider in a data warehousing system. *Communications of the Acm*, 46(11), 111-115.
- Gregg, D. (2009). Developing a collective intelligence application for special education. *Decision Support Systems*, 47(4), 455-465.
- Hart, M., & Porter, G. (2004). The impact of cognitive and other factors on the perceived usefulness of OLAP. *Journal of Computer Information Systems*, 45(1), 47-56.
- Hasan, H., Hyland, P., Dodds, D., & Veeraraghavan, R. (2000). Approaches to the development of multi-dimensional databases: lessons from four case studies. *SIGMIS Database*, 31(3).
- Hasan, K. M. A., Tsuji, T., & Higuchi, K. (2007). An efficient implementation for MOLAP basic data structure and its evaluation. In R. Kotagiri, P. R. Krishna, M.

- Mohania & E. Nantajeewarawat (Eds.), *Advances in Databases: Concepts, Systems and Applications* (Vol. 4443, pp. 288-299). Berlin: Springer-Verlag Berlin.
- John, T. (2008). *Online Analytical Processing*. Unpublished Seminar Report, Cochin University of Science & Technology.
- Kazi, Z., Radulovic, B., Radovanovic, D., & Kazi, L. (2010). *MOLAP Data Warehouse of a Software Products Servicing Call Center*. Paper presented at the MIPRO 2010.
- Korobko, A., & Penkova, T. (2010). On-line analytical processing based on formal concept analysis *Iccs 2010 - International Conference on Computational Science, Proceedings* (Vol. 1, pp. 2305-2311). Amsterdam: Elsevier Science Bv.
- Lehn, R., Lambert, V., & Nachouki, M. P. (1997). *Data warehousing tool's architecture: From multidimensional analysis to data mining*. Los Alamitos: I E E E, Computer Soc Press.
- Lenz, H. J., & Thalheim, B. (2006). OLAP schemata for correct applications. In D. Draheim & G. Weber (Eds.), *Trends in Enterprise Application Architecture* (Vol. 3888, pp. 99-113). Berlin: Springer-Verlag Berlin.
- Lenz, H. J., & Thalheim, B. (2009). A Formal Framework of Aggregation for the OLAP-OLTP Model. *Journal of Universal Computer Science*, 15(1), 273-303.
- McCulley, M. B. (1997). *Effective virtual design of multi-dimensional data models and interfaces*. Paper presented at the Proceedings of the 1997 workshop on New paradigms in information visualization and manipulation.
- MicroStrategy. (1995). The Case for Relational OLAP. Retrieved 14th November 2010 from http://www.microstrategy.com/bi-applications/5Styles/olap_analysis.asp
- MicroStrategy. (2009a). *MicroStrategy Desktop: Advanced Reporting*.
- MicroStrategy. (2009b). *MicroStrategy Desktop: Reporting Essentials*.

- MicroStrategy. (2010a). *MicroStrategy 9 - Basic Reporting Guide*.
- MicroStrategy. (2010b). *MicroStrategy 9 - MDX Cube Reporting Guide*
- MicroStrategy. (2011). OLAP Analysis. Retrieved 17th September 2011 from http://www.microstrategy.com/bi-applications/5Styles/olap_analysis.asp
- Myers, M. (2010). Qualitative Research in Information Systems. *Association for Information Systems*, from <http://www.qual.auckland.ac.nz/>
- Morfonios, K., & Ioannidis, Y. (2008). Supporting the data cube lifecycle: the power of ROLAP. *Vldb Journal*, 17(4), 729-764.
- Morfonios, K., & Ioannidis, Y. (2010). Revisiting the cube lifecycle in the presence of hierarchies. [Article]. *Vldb Journal*, 19(2), 257-282.
- Orbit, G. (2009). Faster, Bigger, Better - Data Integration. <http://it.toolbox.com/blogs/golden-orbit-blog/some-thoughts-after-microstrategy-symposium-34076>
- Pirna, M., & Botezatu, C. P. (2010). *General information on business Intelligence and OLAP systems architecture*.
- Sanford, N. (1970). Whatever Happened to Action Research? *Journal of Social Issues*, 26, 3 - 23.
- SourceMedia. (2011). Information Management. Retrieved 25th August 2011: <http://www.information-management.com/issues/20000701/2349-1.html>
- Tabatabaei, S. H. (2009). *Evaluation of Business Intelligence maturity level In Iranian banking industry*. Lulea University of Technology.
- TechTarget. (2011a). SearchOracle.com. Retrieved 24th August 2011: <http://searchoracle.techtarget.com/definition/multidimensional-database>
- TechTarget. (2011b). SearchOracle.com. Retrieved 28th August 2011: <http://searchdatamanagement.techtarget.com/definition/dimension-table>

- TechTarget. (2011c). SearchOracle.com. Retrieved 28th August 2011:
<http://searchdatamanagement.techtarget.com/definition/snowflaking>
- TechTarget. (2011d). SearchOracle.com. Retrieved 29th August 2011:
<http://searchdatamanagement.techtarget.com/definition/star-schema>
- Venkat. OLTP & OLAP. *Venky's World* from <http://krisvenky.tripod.com/id13.html>
- Westerlund, P. (2008). *Business Intelligence: Multidimensional Data Analysis*.
- Zhou, G., & Chen, H. (2009). *What-if Analysis in MOLAP Environments*. Paper presented at the Proceedings of the 6th international conference on Fuzzy systems and knowledge discovery.